

The international journal of science / 9 March 2023

nature

outlook
Chronic kidney disease

MURKY WATERS

Satellite images reveal rising tide of marine algal blooms

Artificial additives
Is bigger really better when it comes to AI language models?

Turning the tide
Six priorities to mitigate risk of flash floods on drylands

Drying up
Deforestation linked to reduced rainfall in tropical forests

VISIT…

LANZAROTE Caliente.com

Nature.2023.03.11

[Sat, 11 Mar 2023]

- **This Week**
- **News in Focus**
- **Books & Arts**
- **Opinion**
- **Work**
- **Research**
- **Collections**

This Week

- **Celebrate women in science — today, and every day** [08 Mar 2023]
  • International Women’s Day can serve to bring hope, highlight progress, and inspire research communities to continue their efforts to push hard for true gender quality.
- **‘Elite university’ strategies might boost profile and rankings — but at what cost?** [08 Mar 2023]
  • Japan plans to share US$2 billion annually between a handful of top universities. Experience from elsewhere suggests the negatives of this approach might outweigh the positives.
- **Sims-style ‘digital twin’ models can tell us if food systems will weather crises** [07 Mar 2023]
  • From COVID-19 to the war in Ukraine, virtual models could inform global food policy before emergencies unfold.

- EDITORIAL
- 08 March 2023

Celebrate women in science — today, and every day

International Women’s Day can serve to bring hope, highlight progress, and inspire research communities to continue their efforts to push hard for true gender quality.



Nature is asking readers how they’ll be celebrating women in science for this year’s International Women’s Day.Credit: Susana Vera/Reuters

International Women’s Day falls on 8 March; it aims to draw attention to women’s achievements and the fight for gender equality. The day has its critics: too performative, say some; an opportunity for institutions to put on

a facade of change by doing a photo shoot, say others, or to load over-burdened women in their organizations with yet more duties. But they are wrong. There is a need to raise awareness. Women in science still, on average, publish less and win fewer grants and promotions than do men. Harassment, assault and marginalization drive promising researchers out of science, especially those whose race, ethnicity, disability or sexual orientation makes them targets for discrimination.

Activism and action can engender change — if the systems that have long oppressed researchers who are not male can be made to shift. Part of achieving that goal involves raising awareness of what is possible if barriers are broken down. "We don't need another massively shiny campaign," says Jess Wade, a physicist at Imperial College London and a campaigner for gender equality. "We actually need to support the women scientists that we have."

It's in that spirit that *Nature* asked six women researchers how they will be celebrating International Women's Day. We're asking all readers, too — add your thoughts in our poll at the end of this editorial.

Martina Anto-Ocrah is an epidemiologist at the University of Pittsburgh in Pennsylvania who researches sex and gender disparities with an emphasis on women's health and global health, particularly in sub-Saharan Africa. Anto-Ocrah says she wants to celebrate the contribution of social scientists to the advancement of gender equality. Social scientists "are the people who highlight all the cultural issues in our society that hold women back", she says. One example is how, during the COVID-19 pandemic, publication rates for women scientists dropped more markedly than did those for men — confirming that women shouldered a greater share of responsibilities during that time, such as caring for families, leaving less time for research (E. B. Madsen *et al*. *eLife* **11**, e76559; 2022).



Gihan Kamel, Sandra Díaz and Jess WadeCredit: AP/Shutterstock, Carlos Alvarez/Getty, Thomas Angus, Imperial College London

Sandra Díaz is an ecologist at the National University of Córdoba in Argentina and one of the leaders of IPBES, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Díaz wants to celebrate women as science leaders. Although they make up far from half of the researchers running major laboratories or winning big awards, women are increasingly realizing that they can be at the cutting edge of discovery and knowledge production, Díaz notes. “More and more women take leading roles in coming up with groundbreaking ideas, spearheading really risky scientific endeavours, or leading large science-policy bodies.” She says girls are learning that, for a woman, engaging in a scientific career does not necessarily mean working in the shadows as a follower.

Wade will be lauding the success of programmes designed to support early-career researchers, such as the Rising Stars scheme at the Massachusetts Institute of Technology (MIT) in Cambridge, and the many similar schemes it has inspired around the world. Rising Stars offers mentoring and support for researchers from historically marginalized or under-represented groups,

as they move through their careers. The programme is a direct result of the work of MIT biologist Nancy Hopkins, who showed the institute's male leadership that women had less lab space, lower salaries and fewer grants compared with men. This activism led the institute down a path of change, as a new book, *The Exceptions* by Kate Zernike, describes. Wade also writes and edits Wikipedia articles on scientists from historically marginalized groups. She finds that, when she creates or edits a page for a US-based scientist, very often they will have gone through a Rising Stars-style programme.



Tanya Monro, Martina Anto-Ocrah and Aster GebrekirstosCredit: Megan Popelier, Martina Anto-Ocrah, A. Braeunin

Tanya Monro, Australia's chief defence scientist, will be celebrating courage. "It takes courage for women to enter into and persist in the scientific workforce," says Monro, who works in a field that is more male-dominated than most. "It takes courage for women to speak out when they are expected to shoulder the brunt of changing scientific workplaces, so that the girls and women that follow them have better odds of thriving," she adds. "Women and girls in science need every ounce of that courage for

themselves, to overcome the lingering confidence gap many of us carry through life. And I'm glad that these women persist, because it is that courage that allows them to contribute to creating the knowledge and impact that shapes our world."

Gihan Kamel, a physicist at SESAME, the Synchrotron-light for Experimental Science and Applications in the Middle East, based in Jordan, will also be celebrating breaking barriers. "There is progress," she says. "Not least in breaking the extremes in cultural and religious traditions and rules that are made by society and forced on women — usually inherited from one generation to another."

Aster Gebrekirstos, a senior scientist at the World Agroforestry Centre in Nairobi, will be marking women who have succeeded in their roles despite facing significant challenges. These challenges include wars and conflict, says Gebrekirstos, who is from Tigray, a region of Ethiopia that has been at the centre of a devastating conflict. The United Nations Economic Commission for Africa has published *Earth, Oceans and Skies*, an open-access anthology of writing from and about African women scientists (including Gebrekirstos). The book is honest about the hardships women have endured "to reach where we are", Gebrekirstos says. Take a moment to acknowledge those hardships and to advance equity — today, tomorrow and every day after that.

Nature **615**, 187 (2023)

doi: https://doi.org/10.1038/d41586-023-00670-5

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00670-5

- EDITORIAL
- 08 March 2023

‘Elite university’ strategies might boost profile and rankings — but at what cost?

Japan plans to share US$2 billion annually between a handful of top universities. Experience from elsewhere suggests the negatives of this approach might outweigh the positives.



Japan’s government, led by Prime Minister Fumio Kishida (seen here visiting Kobe University) will decide which universities are to receive

excellence funding.Credit: BJ Warnick/Newscom/Alamy

Japan is an exemplar in global research funding. It is one of a small cluster of countries, including Israel, South Korea and Sweden, that consistently spend upwards of 3% of gross domestic product on research and development (R&D). The 2020 average for countries in the Organisation for Economic Co-operation and Development comes in at just under 2.7%.

But investment alone doesn't guarantee R&D success, and the nation's policymakers worry that not enough of Japan's top universities have the kind of autonomy afforded to their counterparts in Western nations, as Nature Index reports in a Japan-focused issue this week. Officials are also concerned about the fact that Japan's universities have been dropping in international rankings.



Nature Index 2023 Japan

The government's response, announced in 2021, is to create a ¥10-trillion (US$75-billion) national endowment fund for research. This will be invested, with the annual returns — expected to be around ¥300 billion — distributed between a select group of institutions. These will be called universities for international research excellence.

Japan's move to separate a subset of its universities is not unprecedented. Indeed, it comes at a time when prominent institutions elsewhere, such as those of the US Ivy League, Australia's Group of Eight and the United Kingdom's Russell Group, are having to think harder about the consequences of their exclusivity, both positive and negative — particularly when it comes to diversity, equity and inclusion.

The inspiration for Japan's fund comes, at least in part, from the endowment model of private institutions in the United States, but there are key differences that must be borne in mind. Harvard University in Cambridge, Massachusetts, for instance, has the world's largest endowment, totalling $50.9 billion at the end of the fiscal year 2022, and this delivers an annual payout to the university of around $2 billion. Some of this is used for research. But Harvard's endowment is made up of contributions from more than 14,000 individual donors, whereas Japan's fund is a public endowment. As Nature Index reports, this is fuelling concerns about possible political interference, which risks undermining the ambition for more autonomy.



[Will Japan's new ¥10-trillion university fund lift research performance?]

Eligible Japanese universities must apply by the end of this month. A board of expert advisers to the nation's Ministry of Education, Culture, Sports, Science and Technology will select the candidates by the autumn. But it is the government, not the research community, that will make the final call, which is far from ideal.

In the United Kingdom, where the practice of concentrating some research funding is well established, it has helped to boost the outputs and profiles of recipient institutions. Members of the Russell Group (comprising 24 research-intensive universities) receive around two-thirds of one of the main sources of public research funding, known as quality-related (QR) funding. This means that the country's roughly 130 remaining eligible universities share just one-third. In 2022, the QR pot exceeded £2 billion (US$2.4 billion).

Reactions to Japan's move have been mixed. Some researchers and policy analysts welcome the country taking a well-trodden path. Others argue that if redistribution of funding is needed, universities in Japan's poorer parts should benefit more, and that this would tap into previously unexplored avenues for boosting economic growth.

It's not often that countries shake up their funding arrangements so radically, but Germany offers an example. In 2005, the nation ushered in its Excellence Initiative, at least in part to distribute more funding among a subset of its largest universities. France has wanted to move in a similar direction, but researchers have resisted the idea.

Predictably, the German initiative has widened the funding gap between those institutions that have received excellence money and those that have not (L. Mergele and F. Winkelmayer *High. Educ. Policy* **35**, 789–807; 2022). And, after more than a decade, the initiative is changing. In its latest iteration, comparatively more funding is going to smaller universities and to interdisciplinary research. The United Kingdom is also facing calls for its next Research Excellence Framework — the scoring system for assessing research and distributing QR funding — to give greater weight to measures of research culture and the quality of a research environment.

Japan's chosen universities must increase the value of their research, develop new disciplines, target global issues and gain worldwide visibility. Loss of visibility is evident among Japan's top universities, some of which have dropped in the rankings in recent years. In 2013, 5 Japanese universities made the top 200 of the Times Higher Education's World University Rankings; a decade later, that figure has dropped to just 2. Concentrating research funding at a handful of institutions is one way to help universities achieve league-table positions, because some ranking methodologies give weight to funding for research, alongside staff numbers and publication data.

Japan is embarking on a path that is likely to enrich a small number of universities and might well have positive results in areas including research outputs and rankings. But amid mounting evidence of rising levels of burnout and staff turnover — and low job satisfaction — in academia worldwide, the country's policymakers should consider both what might be gained and what might be lost by pursuing a strategy of concentrating funding in this way.

Nature **615**, 188 (2023)

doi: https://doi.org/10.1038/d41586-023-00639-4

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00639-4

- WORLD VIEW
- 07 March 2023

Sims-style ‘digital twin’ models can tell us if food systems will weather crises



From COVID-19 to the war in Ukraine, virtual models could inform global food policy before emergencies unfold.

- Zia Mehrabi

In April 2020, I sat outside with a colleague on a deserted campus amid a COVID-19 lockdown. Shelves were emptying in supermarkets across Europe and North America. Panic buying underpinned that, but worker illness and restrictions on mobility and trade were looming: it was not clear how these would create supply-side disruptions. As food-systems scientists, we agreed we simply didn’t have the tools to answer that question.

In recent years, global food security has had shocks from the COVID-19 pandemic, the Russian invasion of Ukraine, extreme weather events and more. My colleagues and I scramble to find convincing answers quickly when development outfits, aid organizations and government think tanks come calling. What would be the impact of COVID-19 mobility restrictions on harvests in sub-Saharan Africa? How would Germany's move away from Russian gas affect global fertilizer production and use? How would heatwaves change the ability of the poorest people to afford food?

As we talked back in 2020, it was clear to both me and my colleague what we needed: next-generation food-systems models, hooked up to real-world data, that consistently capture patterns of food production, transport, processing and consumption, allowing stress-testing and real-time informed responses to systemic shocks.



Food systems: seven priorities to end hunger and protect the planet

Such representations of physical reality — 'digital twins' — already exist in some sectors. They are used in aerospace engineering to prepare for and respond to critical system failures, and in manufacturing to maintain product quality. The European Union is funding an initiative to develop digital twins of Earth systems to tackle climate change and aid protection of nature. Digital twins of food systems will — at least at first — have fewer data to draw on, because the ability to sense food flows varies with location,

food type and stage of processing. But with advances in data generation, computing power and artificial intelligence (AI), I am convinced the time for these models is now.

Historically, major food emergencies such as the global price crises of 1972–75 and 2007–08 have triggered leaps in food-system modelling. But efforts have been fragmented. Economists can create complex models of prices and trade for policy analysis, but with limited geospatial resolution. Agronomists have excellent high-resolution renderings of crop production and yield, but these typically stop at the farm gate. What's missing are the details of how food flows through supply chains — for example, how lentils from Canada make their way to a restaurant in India.

Yet most of what's needed exists. Satellites can monitor crop production and track how yields respond to weather and other factors. Location and traffic data for major roads, railways, waterways and ports are available in many countries. With GPS, we can tell the speed at which goods travel. We know where people live and can quantify their demand for various foods using purchasing-power data, and we have granular information on international trade volumes for major commodities. Logistics researchers have modelled how food flows from producers to consumers for countries including the United States — models just waiting to be draped over representations of real-world infrastructure.

Putting together these pieces will bring immense benefits, allowing near-real-time assessment of the impact of extreme weather, export restrictions or labour shortages. We will be able to model the effects of multiple simultaneous stressors — disease outbreaks, conflicts or energy shocks — and assess how long it would take to re-route flows to alternative ports, offset losses and buffer shocks.

Technical challenges will need attention. Many maps of processing facilities and the agricultural workforce are incomplete, but machine learning and AI offer opportunities to fill these gaps. Integrating public and private data while ensuring producer and consumer privacy is also difficult, but can be done with current cryptographic tools.



Indigenous knowledge is key to sustainable food systems

Implementing digital twins will be an vast undertaking, and must be done right. On average, poorer nations have less capacity to build and use food-system models: they lack funds, they have less-mature monitoring systems and data infrastructure, and more of their economic activity happens in unrecorded informal markets. Global efforts must counter these inequities and democratize access to models.

Promoting data and model sharing will be key. Digital twins built behind closed doors by industry — particularly large commodity traders and the consultants and insurers that serve them — might not best serve the public interest or the resilience of the food system. Hoarding of privileged information is a real concern given increasing corporate power. Terrorism and security risks are considerable, if in-depth understandings of food-system fragilities are made public or inadequately protected from hacking.

Only through partnership between the public and private sectors, strong governance and clear oversight will the risks be mitigated and the full benefits realized. United Nations bodies could guide this effort, bringing together the organizations equipped to maintain the core components and setting rules for data sharing and governance.

To make digital twins of food systems a success, funding agencies must step beyond their silos and pay for initiatives on the same scale as large physics projects.

The world must stop putting off food-systems data science until emergency strikes. With real-time insights, we could see key fragilities in food systems before it's too late.

Nature **615**, 189 (2023)

doi: https://doi.org/10.1038/d41586-023-00640-x

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00640-x

News in Focus

- **One MRI for 4.7 million people: the battle to treat Syria's earthquake survivors** [27 Feb 2023]
  • With only 64 X-ray and 73 kidney-dialysis machines, 7 CT scanners and one MRI machine, doctors in northwest Syria are racing against the clock to treat 8,500 injuries.
- **Asteroid lost 1 million kilograms after collision with DART spacecraft** [01 Mar 2023]
  • Studies reveal final moments before NASA probe crashed into an asteroid.
- **How to stop the bird flu outbreak becoming a pandemic** [01 Mar 2023]
  • From tracking the disease's spread in wild birds to updating human vaccines, there are measures that could help keep avian influenza in check.
- **Ancient genomes show how humans escaped Europe's deep freeze** [01 Mar 2023]
  • A pair of studies offer the most detailed look yet at groups of hunter-gatherers living before, during and after the last ice age.
- **China Initiative's shadow looms large for US scientists** [24 Feb 2023]
  • Anti-Asian scrutiny has only intensified since the controversial programme ended one year ago, researchers say.
- **Data hint at Russia's shifting science collaborations after year of war** [24 Feb 2023]
  • Nature analysis suggests that Russia is increasing partnerships with China and India.
- **In AI, is bigger always better?** [08 Mar 2023]
  • As generative AI models grow larger and more powerful, some scientists advocate for leaner, more energy-efficient systems.
- **Indoor air is full of flu and COVID viruses. Will countries clean it up?** [07 Mar 2023]
  • The current pandemic has focused attention to the importance of healthy indoor air and could spur lasting improvements to the air we breathe.

- NEWS
- 27 February 2023

One MRI for 4.7 million people: the battle to treat Syria's earthquake survivors

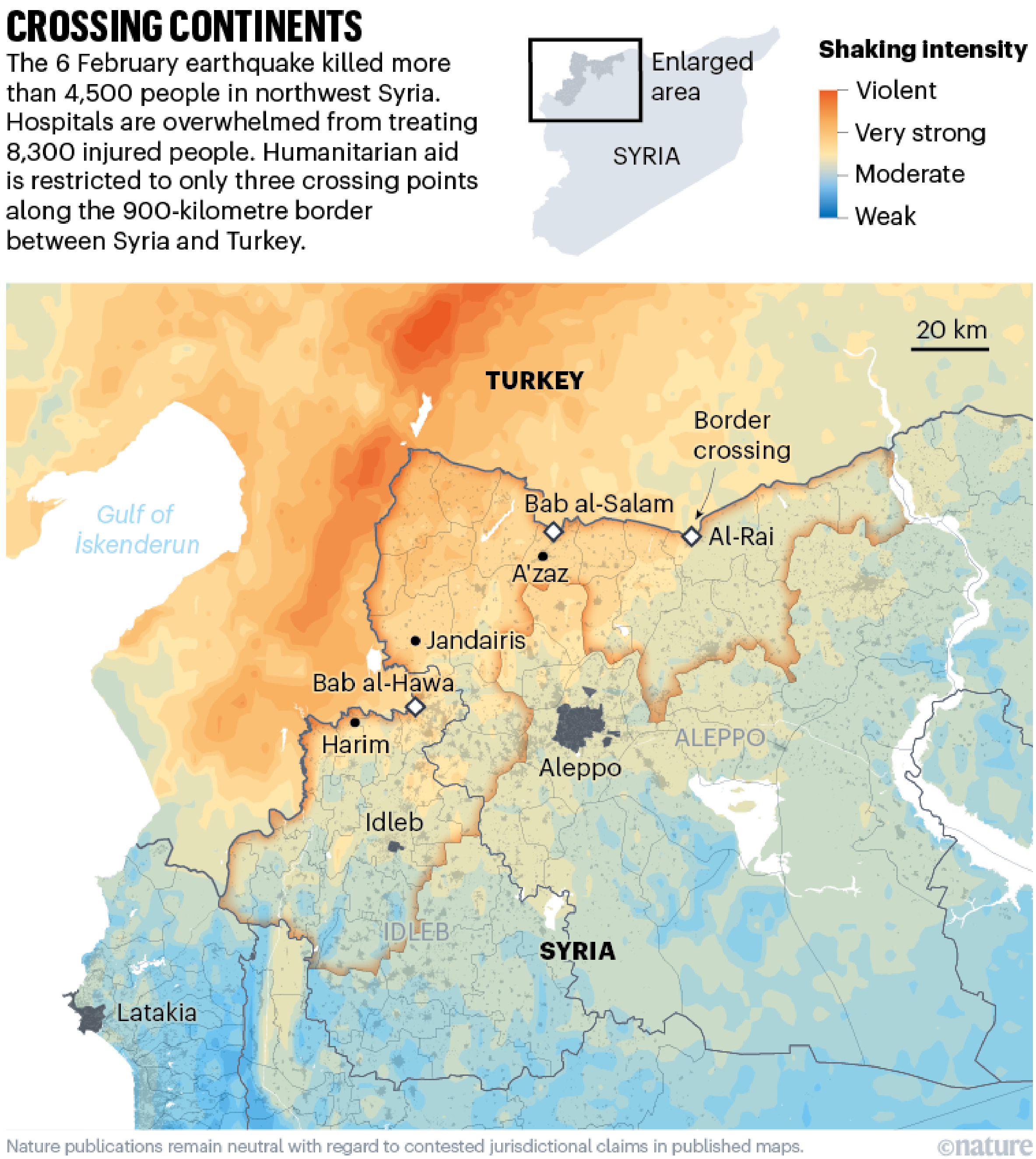
With only 64 X-ray and 73 kidney-dialysis machines, 7 CT scanners and one MRI machine, doctors in northwest Syria are racing against the clock to treat 8,500 injuries.

- Miryam Naddaf



A 7.8-magnitude earthquake struck Turkey and Syria on 6 February, killing more than 47,000 people.Credit: Salwan Georges/The Washington Post via Getty

The 7.8-magnitude earthquake that struck Turkey and Syria on 6 February has killed more than 50,000 people and flattened multiple cities. More than 4,500 of those who died, and 8,500 injured, are in northwest Syria, a region with no unified government that has been cut off from the world for more than 12 years amid a devastating war. Its already-depleted health-care system — 4.7 million people share just one MRI machine — is on its knees. More than 10,000 buildings are completely or partially destroyed, leaving 11,000 people homeless, according to the United Nations Office for the Coordination of Humanitarian Affairs. The earthquake also destroyed warehouses that store medicines. *Nature* spoke to doctors, engineers and other experts on the ground, as well as those helping remotely from Europe and elsewhere in the Middle East. This is what they have said.



Source: Earthquake intensity data, USGS; Waterways, Humanitarian OpenStreetMap Team; Buildup, AtlasAI; Administrative boundaries, UN Office for the Coordination of Humanitarian Affairs.

After the 2011 Arab Spring, Syria's government used military force against all opposition. Around 60% of northwest Syria's 4.7 million people are internally displaced, having fled bombing attacks by the government of Bashar al-Assad with military backing from Russia.

The UN and aid agencies are struggling to get supplies and expertise to earthquake-affected areas, because there are only three temporary crossing points along the 911-kilometre border with Turkey (see ‘Crossing continents’).

Medical emergency

Hospitals in this region have been overwhelmed as they attempt to help thousands of injured people in spaces with severely limited beds, medical supplies, surgical equipment and intensive-care facilities.

More than 8,500 injured people need to be accommodated in only 66 functional hospitals, which are providing 1,245 beds for short hospital stays, according to the World Health Organization (WHO). Moreover, all of northwest Syria has just 86 orthopaedic surgeons, 64 X-ray machines, 7 computerized tomography (CT) scanners and one magnetic resonance imaging (MRI) machine across the region, according to WHO-compiled data published towards the end of 2022. Gynaecologist Ikram Haboush, director of the sole public maternity hospital in the city of Idleb, says “the medical situation in northwest Syria is catastrophic”.

“The supply of antibiotics ran out from day three” after the earthquake, says Abdulkarim Ekzayez, an epidemiologist at King’s College London, who now fears widespread infections. “We have used the medications and serums that would have lasted us for four to six months in two to three days,” adds Haboush. The WHO has started airlifting medicines and medical supplies, but says northwest Syria also needs essential diagnostic equipment such as X-ray machines.

The region’s critical lack of health care is partly caused by hospitals and medical staff being targeted during the war, explains Ekzayez, who is a co-investigator on a UK-funded project called Research for Health System Strengthening in Syria. In a separate study, Ekzayez estimated that as of June 2021, 350 medical facilities had been attacked and 930 health-care personnel killed.

“People are running all over the place to make use of any existing resources, including basic ambulances,” Ekzayez says. “The medical staff have been working non-stop,” adds Haboush, who was working at the maternity hospital in Idleb when the first earthquake hit at 01:17 universal time. The hospital occupies the fifth and sixth floors of a building. “We had to evacuate the incubators and move all the babies to the ground floor,” Haboush recalls.

According to three doctors who *Nature* spoke to in northwest Syria, the most common injuries are limb fractures, trauma injuries, crush injuries including ‘crush syndrome’ (damages that result in organ dysfunction including kidney failure) and bleeding.

People with crush syndrome need intensive care and dialysis, says cardiologist Jawad Abu Hatab, who is dean of medicine at Free Aleppo University in northwest Syria. However, the region has only 73 renal-dialysis machines, according to the WHO-compiled data.

“There were also cases of cardiac arrests from the shock and horror of the catastrophe,” Abu Hatab adds. The medical community is also preparing to deal with more cases of post-earthquake trauma, especially among children and women.

Doctors say that they urgently need more dialysis machines and orthopaedic-surgery equipment, along with painkillers and antibiotics. “These supplies were not abundant before the earthquake” and will now run out, says Haboush.



Rescuers searched for survivors in the rubble of collapsed buildings in Harim, one of the towns with the highest number of deaths and injuries.Credit: Omar Haj Kadour/AFP via Getty

Naked-eye damage assessments

Volunteer engineers are going house-to-house checking which of some 8,500 earthquake-damaged buildings are safe enough for people to return to. But lacking in the necessary tools and safety equipment, they are resorting to tapping walls with household implements such as simple hammers and making decisions using the naked eye.

It is painstaking work. As of 25 February, around 2,644 of the damaged buildings had been assessed, according to the Association of Free Syrian Engineers , a voluntary group based in A’zaz, near Aleppo, that is coordinating the effort.

At the same time, experts from the Syrian diaspora are helping with virtual assessments in places inaccessible to local engineers. Residents of damaged

buildings are taking photos and videos of the interiors and sending them to members of the Syrian Engineers Association in Qatar, based in Doha.

“Assessing the structural status of buildings is as urgent as the medical situation and should not be underestimated because aftershocks are still ongoing, says Muhammad Azmie Tawackol, a civil engineer and the association’s president. But in northwest Syria, people are having to use whatever methods are available, he adds.

The engineers are trying to estimate whether buildings are inhabitable (safe with minor cracks), temporarily unusable (needing reinforcement) or unsafe — in which case occupants must evacuate immediately.

Buildings in danger of collapsing are being reinforced with whatever materials are available, irrespective of whether they are suitable. Carbon-fibre-reinforced polymers would work better for seismic reinforcement, Mohammad Khear Hayek, an engineer with the volunteers association, told *Nature*. Instead, they are having to use brittle industrial iron. “We are in an emergency situation, so we must respond quickly using the resources that we have,” Hayek says.

The engineers association based in northwest Syria is collating data daily and “the next step is to analyse the reports and produce statistical studies, which will be important in the reconstruction phase”, says association member Ali Hallak, a computer engineer.

Many buildings will need to be rebuilt, but there is a shortage of engineers in northwest Syria, says Hayek. Before the earthquake, “we talked about a need to train engineers on reconstruction according to appropriate safety standards”, Hayek says. “Now this has become a necessity,” he adds.

Nature **615**, 193-194 (2023)

doi: https://doi.org/10.1038/d41586-023-00547-7

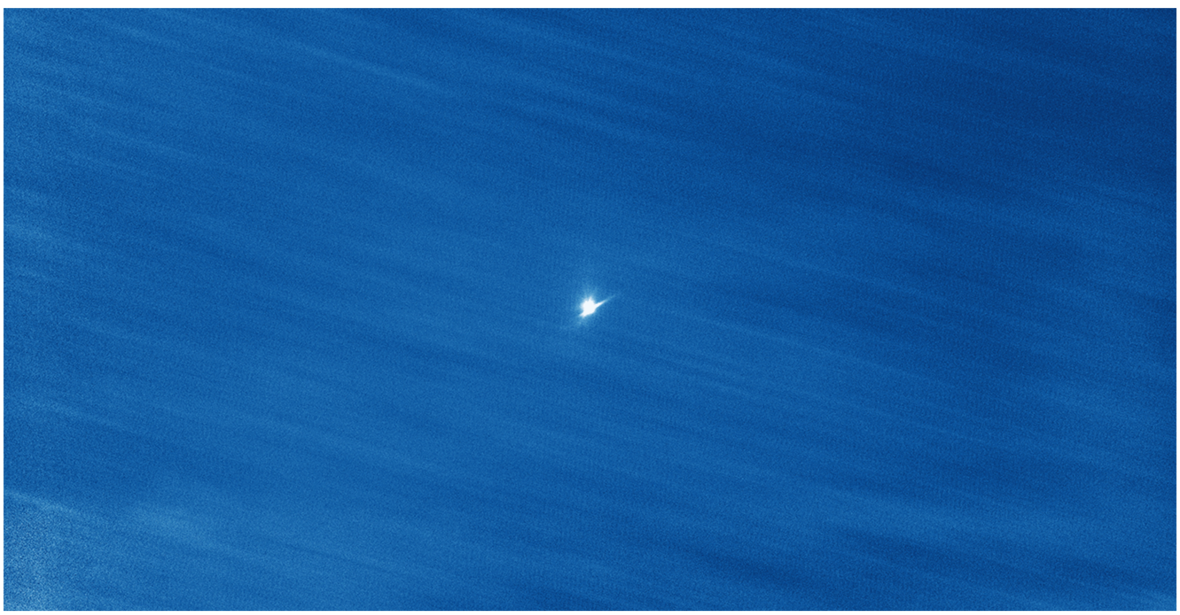
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00547-7

- NEWS
- 01 March 2023
- Clarification 02 March 2023

Asteroid lost 1 million kilograms after collision with DART spacecraft

Studies reveal final moments before NASA probe crashed into an asteroid.

- Alexandra Witze



Images from the first month after DART’s asteroid collision. Each frame is the average over an entire night’s observations. The Didymos asteroid system is in the centre and stars appear as streaks.Credit: University of Canterbury Ōtehīwai Mt John Observatory/UCNZ

Last September, NASA's Double Asteroid Redirection Test (DART) spacecraft smashed into an asteroid, deliberately altering the rock's trajectory through space in a first test of planetary defence. Now scientists have deconstructed the collision and its aftermath — and learnt just how successful humanity's punch at the cosmos really was.

DART, which was the size of a golf cart, collided with a Great Pyramid-sized asteroid called Dimorphos. The impact caused the asteroid's orbit around another space rock to shrink — Dimorphos now completes an orbit 33 minutes faster than before the impact, researchers report[1] today in *Nature*. This means that if a dangerous asteroid were ever detected heading for Earth, a mission to smash into it would probably be able to divert it away from the planet.

DART's success has been reported before; now, five studies in *Nature* describe the final moments of the crash and how it affected the asteroid. One group combined data on the spacecraft's trajectory with photographs of the asteroid's surface just before impact[2]. As DART hurtled towards Dimorphos at more than 6 kilometres per second, the first part that hit was one of its solar panels, which smashed into a 6.5-metre-wide boulder. Microseconds later, the main body of the spacecraft collided with the rocky surface next to the boulder — and the US$330-million DART shattered to bits.

The impact ejected at least one million kilograms of rock from Dimorphos's 4.3-billion-kilogram mass. The debris formed a tail that stretched for tens of thousands of kilometres behind the asteroid. Various telescopes watched over weeks as the tail shifted and evolved under the pressure of the Sun's rays; the Hubble Space Telescope even detected a second tail, which had disappeared by 18 days after the impact[3].

Dimorphos is 151 metres wide and orbits the larger asteroid Didymos. NASA's goal was to alter Dimorphos's orbit enough for astronomers to spot the changes by monitoring the brightness of the pair over time using ground-based telescopes. Neither asteroid is, or will ever be, a threat to Earth.

Images taken as DART approached Dimorphos on 26 September show the asteroid looking like an egg covered in boulders. It seems to be a loose pile

of rubble barely held together by gravity — whose surface would probably shatter spectacularly when DART hit it.

Crash test

The discovery of more details is helping researchers to understand why the impact was so successful in shunting the asteroid, says Carolyn Ernst, a planetary scientist at the Johns Hopkins University Applied Physics Laboratory in Laurel, Maryland.

One factor is that the spacecraft hit a spot around 25 metres from the asteroid's centre, maximizing the force of its impact. Another is that large amounts of the asteroid's rubble flew outwards from the impact. The recoil from this force pushed the asteroid further off its previous trajectory. Researchers estimate that this spray of rubble meant Dimorphos' added momentum was almost four times that imparted by DART[4].

Although NASA has demonstrated this technique on only one asteroid, the results could be broadly applicable to future hazards, researchers say. "It means that we can quickly design a mission to deflect an asteroid if there is a threat, and we know that this has a very high chance of being effective," says Franck Marchis at the SETI Institute in Mountain View, California, who is also chief scientific officer at the telescope manufacturer Unistellar in Marseille.

"If you had asked me 30 years ago, 'Can we be confident we won't be wiped out by a giant killer asteroid a week from next Tuesday?' I would have had to say no," adds Tom Statler, DART's programme scientist at NASA headquarters in Washington DC. Now that astronomers have surveyed the skies to identify nearly all the dangerous asteroids — and now that DART has been shown to work —"we will know what to do about it when something new is found", he says.

Researchers are continuing to work through the DART data to learn more about the physics, chemistry and geology of both Dimorphos and Didymos. This work is being done with the help of a network of amateur astronomers co-led by Marchis. The network's members observed the asteroids with their

telescopes before, during and after the impact. They discovered that the rocks seemed to become significantly redder immediately after the spacecraft hit[5].

A later colour change to blue was spotted by NASA's Infrared Telescope Facility in Hawaii, as reported by Cristina Thomas, a planetary scientist at Northern Arizona University in Flagstaff, at a December meeting of the American Geophysical Union. "We think this is likely because we have a lot of material from Dimorphos thrown off," she says. The impact blasted through the asteroid's weathered interior and exposed part of its insides, making the asteroid appear temporarily blue, a few hours after impact.

Scientists will get a close-up view of the impact's aftermath in late 2026, when a European Space Agency craft named Hera will arrive at Dimorphos. Researchers are keen to see what sort of impact crater the DART smash left behind. In the meantime, detailed scientific observations will continue for several weeks, until the movement of Didymos and Dimorphos away from Earth makes them too faint for many telescopes to see, and then takes them behind the Sun.

"It's really thrilling to still be watching the evolution of what's happening," says Thomas.

Nature **615**, 195 (2023)

doi: https://doi.org/10.1038/d41586-023-00601-4

Updates & Corrections

- **Clarification 02 March 2023**: This story has been updated to clarify that the asteroid changed colour from red to blue after impact. Frank Marchis's affiliation has also been amended.

References

1. Thomas, C. A. *et al.* *Nature* https://doi.org/10.1038/s41586-023-05805-2 (2023).

2. Daly, R. T. *et al. Nature* https://doi.org/10.1038/s41586-023-05810-5 (2023).

3. Li, J.-Y. *et al. Nature* https://doi.org/10.1038/s41586-023-05811-4 (2023).

4. Cheng, A. F. *et al. Nature* https://doi.org/10.1038/s41586-023-05878-z (2023).

5. Graykowski, A. *et al. Nature* https://doi.org/10.1038/s41586-023-05852-9 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00601-4

- NEWS EXPLAINER
- 01 March 2023

How to stop the bird flu outbreak becoming a pandemic

From tracking the disease's spread in wild birds to updating human vaccines, there are measures that could help keep avian influenza in check.

- Saima May Sidik



A goose being vaccinated against avian influenza in China.Credit: Wei Liang/China News Service/Getty

Fears are rising about bird flu's potential to spark a human pandemic, as well as its destruction of wildlife and farmed birds. An 11-year-old girl tragically died in Cambodia last week after catching avian influenza. That followed reports earlier this year of the virus spreading from mammal to mammal through a mink farm, and causing mass mortality in Peruvian birds and sea lions. Since the beginning of 2022, more than 50 million poultry birds in the United States, and a similar number in Europe, have either died of the disease or been killed in efforts to stem its spread. Can bird flu be stopped, and if yes, how?

Protecting poultry

Poultry farms are a key battleground in the fight against H5N1, the strain of avian influenza that is currently circulating. Outbreaks on farms threaten food security and provide opportunities for the virus to spread to farm workers. For decades, farmers have controlled the disease by culling infected animals. But now, with many countries experiencing outbreaks on dozens of farms every month, this is becoming untenable.

Some countries, including China, vaccinate poultry to limit the spread and severity of bird flu, and other governments around the world are now implementing vaccination policies or contemplating doing so. One problem with existing vaccines is that they cause birds to test positive for the virus, meaning farmers can't guarantee their birds are free of H5N1. This has "huge international trade and export implications", says Keith Poulsen, a specialist in infectious diseases who directs the Wisconsin Veterinary Diagnostic Laboratory in Madison.

Scientists are in the early stages of developing vaccines that might solve this problem. Microbiologist Adel Talaat at the University of Wisconsin–Madison and his colleagues are developing a vaccine that uses only a small part of the virus's DNA. Tests targeting other genetic regions could differentiate between birds that have been vaccinated and those that are infected.

Poultry farmers could also raise a wider variety of bird breeds in order to stop the virus, says Nichola Hill, an ecologist at the University of

Massachusetts in Boston. In Asia, where farmers have a long history of navigating outbreaks of bird flu, some have switched to breeds that are less susceptible to the virus.

Conserving wildlife

H5N1 has become entrenched in wild bird populations over the past year, but there are “some small Band-Aids we can put on things”, says epidemiologist David Stallknecht at the University of Georgia in Athens. Administering vaccines to wild birds is logistically difficult. So, for the most part, birds have to develop resistance to the disease through being infected, and many will die in that process.

Vaccines could help to protect certain species, Stallknecht says. Bald eagles (*Haliaeetus leucocephalus*), for example, can be severely affected by the virus, and some scientists are worried about the long-term impact of bird flu on the population. But the strategy could only be used for species under grave threat when “you’re doing everything you possibly can to keep them on the planet”, he says.

At the moment, Stallknecht and other wildlife researchers are trying to understand which species of wild bird are most severely affected by avian flu, and the implications this has for the spread of the disease. As well as helping scientists target conservation measures, this research could give farmers a better idea of when bird flu might be heading their way if, for example, it is combined with when certain birds are known to migrate.

That knowledge could help farmers target measures to protect poultry, such as cleaning up grain that could attract wild birds and washing boots before entering farms. “It’s extremely hard to do that 365 days of the year,” Hill says. Shorter time frames are more feasible.

Stopping a human pandemic

The death of the girl in Cambodia — and the fact that her father also tested positive for avian influenza — has renewed concerns about whether bird flu could spark widespread infection in people, or even a pandemic. “That’s

hard to say," says Thijs Kuiken, a veterinary pathologist at the Erasmus University Medical Center in Rotterdam, the Netherlands.

Ancestral versions of today's H5N1 virus have been circulating among birds for about 25 years and have not yet gained the ability to spread between humans. That leads Kuiken to think that the risk of a human pandemic is low. But the recent uptick in cases among wild birds, and the discovery that the virus can be transmitted between mammals, increase the risk that the virus could start spreading in humans. Kuiken would like to see increased surveillance of people who work in the poultry sector to make sure anyone who becomes infected is quickly detected and isolated.

If bird flu does trigger a human pandemic, there are a number of tools for combating the disease. Approved human vaccines against avian flu exist, and the World Health Organization monitors the evolution of H5N1 so that these vaccines can be updated appropriately. In the United States, the Biomedical Advanced Research and Development Authority has a stockpile of vaccines, although the supply is too low to be used to vaccinate the world. Animal studies and observational data in humans suggest the antiviral drug Tamiflu is effective against H5N1 in people[1], although there have been reports of resistant strains[2]. Non-pharmaceutical tools including face masks can also limit disease spread.

For a world still reeling from COVID-19, the prospect of another pandemic is alarming. Bird flu's current mortality rate in humans is around 50%, although that would be likely to drop if the virus gained the ability to infect cells in the upper respiratory tract — a prerequisite for efficient human-to-human spread. But several scientists say that an H5N1 pandemic would probably be more manageable than COVID-19 because of the drugs and vaccines that are already available, and because of tools such as mRNA vaccines that were developed as a result of COVID-19. "Not to say it won't be a mess," says Stallknecht, "but it probably won't be as bad as it could be."

Hill agrees that humanity has the tools needed to keep the virus in check. "The question is control at this point, and preventing a human pandemic," she says. "And I think both of those are achievable goals."

Nature **615**, 196-197 (2023)

doi: https://doi.org/10.1038/d41586-023-00591-3

References

1. Smith, J. R. *J. Antimicrob. Chemother.* **65**, ii25–ii33 (2010).
2. de Jong, M. D. *et al. N. Engl. J. Med.* **353**, 2667–2672 (2005).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00591-3

- NEWS
- 01 March 2023

Ancient genomes show how humans escaped Europe’s deep freeze

A pair of studies offer the most detailed look yet at groups of hunter-gatherers living before, during and after the last ice age.

- Ewen Callaway



These 14,000-year-old skulls were found in western Germany. Their genetic ancestry suggests that human populations migrated in response to Europe’s

changing climate.Credit: Jürgen Vogel/LVR-LandesMuseum Bonn

Like retirees who flock to the Costa del Sol, ancient European hunter-gatherers sought out Spain's warmer climate during the peak of the last ice age.

A pair of ancient-genome studies shows that humans who had holed up in the Iberian Peninsula repopulated Western Europe after the retreat of glaciers that covered large parts of the continent from about 26,000 to 19,000 years ago[1,2].



Ancient African genomes offer glimpse into early human history

The research — based on newly sequenced genomes of more than 100 individuals — offers the most detailed look yet at groups of hunter-gatherers who lived in Europe before, during and after the last ice age.

Ancient genomes from this period are scarce, so "even adding one data point is really important", says Mateja Hajdinjak, a molecular biologist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, who was not involved in either study.

Ancestral traces

Homo sapiens migrating out of Africa reached Europe at least 45,000 years ago — and possibly earlier. But the handful of ancient genomes from this period suggest that these pioneers left no genetic trace in later hunter-gatherers.

Instead, a landmark 2016 study[3] identified a genetic signature in 35,000-year-old remains from the Goyet cave system in Belgium, which persisted in hunter-gatherer populations that lived tens of thousands of years later. The Goyet remains were associated with Aurignacian artefacts, a Europe-wide material culture known for its elaborate cave-wall art and 'Venus' figurines.

But the 2016 study raised a mystery. The Goyet ancestry was missing in remains from a roughly 20,000-year period leading up to and during the peak of the ice age, before reappearing later in hunter-gatherers in western Europe. "Where were these people hiding for 20,000 years?" asks Cosimo Posth, a palaeogeneticist at the University of Tübingen in Germany.



Ancient genomics is recasting the story of the Americas' first residents

To find out, Posth and his colleagues sequenced ancient DNA from 116 hunter-gatherers who lived in Europe and western Asia between 35,000 and 5,000 years ago, and analysed previously sequenced genome data from hundreds more.

Among the trove they detail in *Nature* on 1 March[1] are several pre-ice-age individuals from sites in France and Spain. Their remains were found with artefacts attributed to another pan-European material culture known for its

figurines, called the Gravettian. When Posth's team looked at these people's genomes, the researchers found a direct link to the earlier Goyet population that had seemingly vanished.

The researchers traced this ancestry to 21,000- and 23,000-year-old individuals from sites in northern Spain and southwestern France, respectively. These sites are linked to another culture called the Solutrean.

A separate study reported on 1 March in *Nature Ecology and Evolution*[2] found that a 23,000-year-old man from a site in southern Spain — also linked to Solutrean artefacts — also had Goyet ancestry.

The two studies suggest that the Iberian Peninsula was a refuge for hunter-gatherers as the climate cooled and glaciers ensconced northern Europe. The genetic signature — the same one found in Goyet 35,000 years ago — later pops up across western Europe and even into Poland after Europe's climate warmed.

Ice-age refuges

Iberia wasn't Europe's only ice-age holdout, says Posth. His team discovered a genetic signature in post-ice-age people in northern Italy that eventually reached Sicily. But these people were not related to humans who were in Italy before the peak of the Ice Age, nor to the western hunter-gatherer groups.



[Oldest DNA from a Homo sapiens reveals surprisingly recent Neanderthal ancestry]

Instead, they descend from people who saw out the coldest periods in southeastern Europe, probably in the Balkans, says Posth. There is currently no ancient DNA from this period to confirm that hunch, he adds, leaving a “big empty spot on the map”.

Colin Wren, an archaeologist at the University of Colorado Colorado Springs, says those findings confirm his own work, suggesting that ice-age Italy was less hospitable to humans than was Iberia[4]. Italy was once connected to Croatia by a now-submerged plain, so it makes sense that eastern hunter-gatherers moved in, he adds.

The studies paint a dynamic picture of European hunter-gatherer populations, which don’t always line up with cultural shifts, says Natasha Reynolds, a palaeolithic archaeologist at the University of Bordeaux, France. They also show that, for parts of Europe at least, the coldest period of the ice age wasn’t as inhospitable as it’s often made out to be. “People weren’t just huddling in caves waiting for the glaciers to retreat,” she says.

Nature **615**, 197-198 (2023)

doi: https://doi.org/10.1038/d41586-023-00611-2

Read the related News & Views: The genomic history of ice-age Europeans.

References

1. Posth, C. *et al. Nature* **615**, 117–126 (2023).
2. Villalba-Mouco, V. *et al. Nature Ecol. Evol*. https://doi.org/10.1038/s41559-023-01987-0 (2023).
3. Fu, Q. *et al. Nature* **534**, 200–205 (2016).
4. Wren, C. D. & Burke, A. *PLOS ONE* **14**, e0217996 (2019).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00611-2

- NEWS
- 24 February 2023
- Update 27 February 2023

China Initiative’s shadow looms large for US scientists

Anti-Asian scrutiny has only intensified since the controversial programme ended one year ago, researchers say.

- Natasha Gilbert



Activists in Los Angeles, California, protest against anti-Asian racism, which some say that the China Initiative helped to perpetuate.Credit: David McNew/AFP/Getty

One year after the US government ended its controversial China Initiative, scientists of Chinese heritage say that they are still being targeted unfairly and fear for their safety.

The initiative — which was aimed at safeguarding US laboratories and businesses from espionage — created the perception of bias against researchers of Chinese descent, said assistant attorney-general Matthew Olsen when shutting it down in February 2022, although he denied that the programme had actually used racial profiling. While it was active, more than 150 people were criminally charged for actions such as failing to disclose funding or partnerships with institutions in China, according to an analysis by *MIT Technology Review*. Nearly 90% of them were of Chinese heritage. Many of the charges brought by the US Department of Justice (DoJ) after the initiative's launch in 2018 were eventually dropped or dismissed, and some prosecutions ended in acquittal.



The controversial China Initiative is ending — researchers are relieved

The climate of fear and anxiety hasn't gone away — researchers are just being pressured in a new way, says Jenny Lee, a social scientist at the University of Arizona in Tucson who studies research collaborations and geopolitics. Since the initiative's official shutdown, the US government has adopted various anti-China policies. And although the DoJ is pursuing fewer criminal charges, it says that it will work increasingly with federal agencies

to investigate researchers and issue civil and administrative penalties for noncompliance. Universities are also taking a more active role in assisting investigations and pursuing potential wrongdoing, sources tell *Nature*.

"I'm sorry to say that it has only intensified," says Gang Chen, a mechanical engineer at the Massachusetts Institute of Technology in Cambridge, who was arrested in January 2021 under the China Initiative, only for the DoJ to drop the charges a year later. He and others who have had their lives upended by the initiative have been speaking out about the damage that it has done.

"The government has not done enough" to ease the situation, Chen adds. The DoJ did not respond to *Nature*'s request for comment.

Policy shift

One example of a university taking a more active role in the initiative's wake was reported by *The San Diego Union-Tribune* in December 2022. Xiang-Dong Fu, a molecular biologist at the University of California, San Diego (UCSD), was forced to quit his position after the university accused him of hiding ties to China. UCSD said he had violated its conflict-of-commitment policy by accepting travel reimbursements from Chinese institutions that he had visited, and had failed to disclose Chinese grants that bore his name. Fu denies any wrongdoing, according to the *Tribune*.



The US Department of Justice dropped charges against mechanical engineer Gang Chen in early 2022. He was arrested one year earlier for hiding ties to China.Credit: Tony Luong/New York Times/Redux/eyevine

Universities reject the idea that they are unfairly targeting researchers of Chinese heritage. According to Toby Smith, vice-president for science policy and global affairs at the Association of American Universities (AAU) in Washington DC, US institutions acknowledge the considerable research contributions from these scientists. Universities are working to ensure that all faculty members are disclosing information properly, he adds.

But he calls on US funding agencies to provide greater clarity for universities on what counts as an offence and what are appropriate and fair sanctions.

Scientists need support, says Gisela Kusakawa, executive director of the Asian American Scholar Forum, a non-profit organization based in New York City. Universities and agencies should provide training for scientists on

how to complete disclosure forms, and they must allow scientists the opportunity to revise completed forms to ensure that they are correct, she says.

The illusion of improvement

In the past year, the US government has adopted several policies and positions that have perpetuated the narrative that scientists from China are potential spies, Lee says. In August 2022, the US Congress passed into law the CHIPS and Science Act, which earmarks an extra US$280 billion for research and innovation and includes measures designed to tighten research security. For example, it asks US institutions to report gifts of $50,000 or more from a foreign government, down from the previous minimum of $250,000.



Scientists' fears of racial bias surge amid US crackdown on China ties

In January, Congress also voted to form a bipartisan committee to assess the economic and competitive threats that China poses to the United States. The AAU has said that the creation of the committee signals an intent in Congress to monitor China's influence on the nation's scientific enterprise.

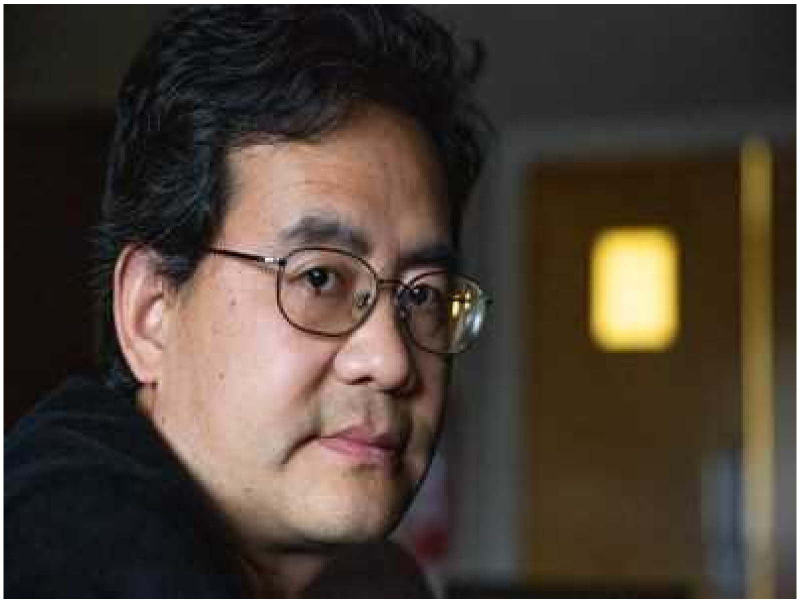
The US government has caught genuine Chinese spies stealing trade secrets and scientific and technological developments. But many say that the government's broad-brush scrutiny of researchers of Chinese descent is excessive, and could actually harm national security by demonizing scientific collaboration with China and driving out scientists who contribute to US scientific prowess. In a survey published this week[1], Lee says, she found a link between fears of racial profiling and a desire among scientists to return to China.

The end of the China Initiative gave the illusion that researchers of Chinese heritage would be targeted less, she says, but "the chilling effect" is "still very much at play".

Afraid of doing research

Researchers unjustly accused under the China Initiative and now rebuilding their lives and careers are emblematic of this situation.

Xiaoxing Xi, a physicist at Temple University in Philadelphia, Pennsylvania, was arrested at gunpoint in front of his family by the DoJ in 2015. Although this was before the initiative launched officially in 2018, scrutiny of researchers of Chinese heritage had begun years earlier. Xi was accused of passing information to scientists in China about restricted technology. The DoJ eventually dropped the charges.



[‘I lost two years of my life’: US scientist falsely accused of hiding ties to China speaks out]

Xi has been seeking damages for harm he experienced as a result of his arrest, and is appealing against a ruling in March 2022 that dismissed his claims. He is nervous about applying for federal research funding, and spends much of his time following the cases of targeted scientists and giving talks to raise awareness about anti-Asian sentiment. Before his arrest, he was juggling 9 research projects and had 15 people working in his lab. Now, he is working on just one project and has one researcher on his team.

“I am afraid of doing any research,” he says. “We always live in fear.”

Chen is similarly afraid to apply for federal research funding, concerned that the government could misuse the forms against him as they did before, he says. To feel more secure, he has switched from researching nanotechnologies with obvious commercial applications to doing more-fundamental science, exploring the solar evaporation of water. He also rarely answers e-mails from researchers or students in China who write asking questions about his research papers.

Anming Hu, a nanotechnology researcher at the University of Tennessee in Knoxville, who was indicted for hiding ties with China in 2020 and put under house arrest for more than a year before being acquitted, is also trying

to get his research back on track. He has spent the past year rebuilding his lab, but has had trouble securing any funding. Currently, he has two graduate students on his team; before his arrest he had six, he says, adding that he won't now take on students or researchers from China because it's too risky.

"If nothing had happened to me, I would have climbed to a much higher research level," he says. "But I do my best."

Nature **615**, 198-199 (2023)

doi: https://doi.org/10.1038/d41586-023-00543-x

Updates & Corrections

- **Update 27 February 2023**: This story was updated to include a link to Jenny Lee's newly published study on geopolitics interfering with global science.

References

1. Lee, J. J. & Li, X. *Rev. High. Educ.* **46**, 285–309 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00543-x

- NEWS
- 24 February 2023

Data hint at Russia’s shifting science collaborations after year of war

Nature analysis suggests that Russia is increasing partnerships with China and India.

- Richard Van Noorden



University buildings in Kharkiv, Ukraine, were damaged in a Russian missile strike earlier this month.Credit: Viacheslav Mavrychev/Suspilne Ukraine/JSC "UA:PBC"/Global Images Ukraine/Getty

A year into Russia's war on Ukraine, its effects on global research collaborations might be starting to show up in the scientific literature.

Nature analysed co-authorship patterns on papers in the Scopus database. The results suggest that, in 2022, an increased share of Russia's internationally collaborative papers had co-authors from China and India, whereas the proportion co-written with US or German authors fell. Ukraine, meanwhile, has sharply reduced its scholarly ties with Russia, and seems to have increased research connections with Poland.

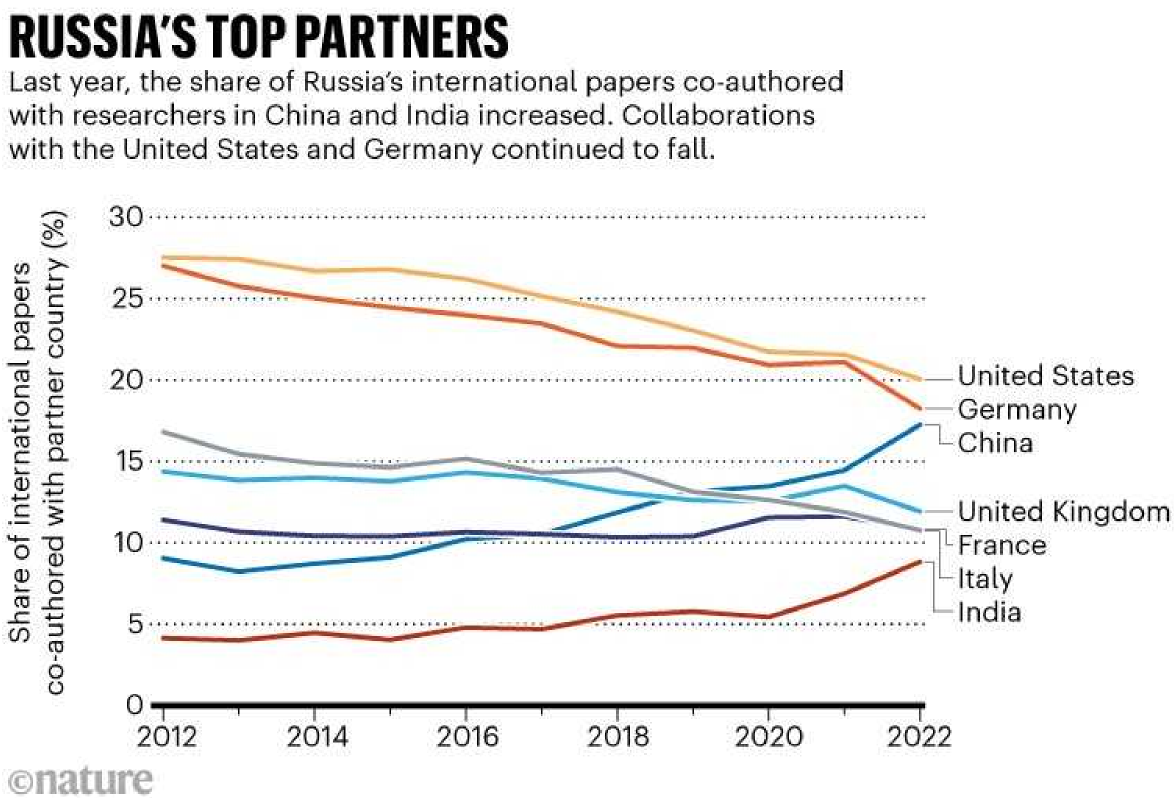


The fight to keep Ukrainian science alive through a year of war

The data available so far can only hint at possible changes, which will become more apparent later this year. Many of the papers published in 2022 were submitted to journals well before Western institutions halted scientific partnerships with Russia in response to the full-scale invasion, which began on 24 February. And databases of scientific papers will continue filling up their 2022 collections for another month or two, owing to routine indexing delays, so the raw numbers of papers will continue to increase. (Relative proportions of co-authorship are unlikely to vary at this stage.)

Co-authorship changes

Last year, just over a quarter of Russia's papers and review articles — as recorded in Scopus — were internationally co-authored, a similar proportion to the year before. China is now close to overtaking the United States and Germany to become Russia's top research partner, and India has soared up the rankings to seventh place (see 'Russia's top partners'). Among Russia's top 25 collaborators, the only other countries that have increased their share of its international co-authorships in 2022 are Kazakhstan and Iran.

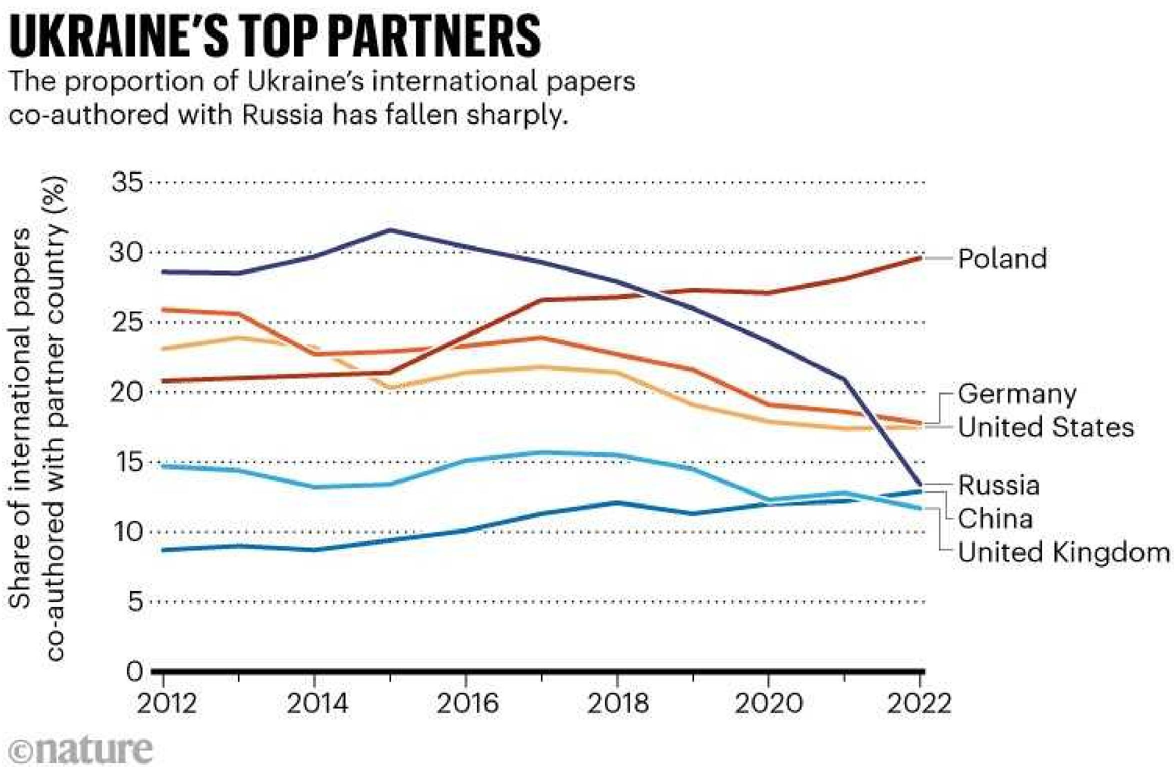


Source: *Nature* analysis of Scopus data.

Kieron Flanagan, a science-policy researcher at the University of Manchester, UK, cautions that co-authorships are an imperfect proxy for real research-collaboration patterns, and that it's too early to draw conclusions from the available data. But he says that it does look like Russia has seen a long-term relative decline in research collaborations with Western countries, and this trend could have accelerated in 2022.

China is increasing its scientific co-authorship share with most countries except the United States, so its rising influence on Russia might not be particularly unusual, says Caroline Wagner, a science and policy researcher at the Ohio State University in Columbus. “The Russia–Ukraine conflict does not appear to have interrupted Russia and China’s cooperation,” she says.

In Ukraine, meanwhile, the proportion of papers with Russian co-authors fell sharply in 2022, and Poland is the country’s clear leading research partner (see 'Ukraine’s top partners’). Around 38% of Ukraine’s papers in 2022 were internationally co-authored — a slightly higher proportion than in previous years.



Source: *Nature* analysis of Scopus data.

Nature found similar patterns in a parallel analysis of scientific papers in a different scholarly database, Dimensions (which differs from Scopus in its choice of journals to index). But according to the Dimensions data,

Ukraine's collaboration with Poland has not increased: it has remained at around the same level since 2021.

Cutting research ties

Many research organizations cut collaboration activities with Russia shortly after its invasion of Ukraine. Some funders, including those in Poland and Germany, strongly discouraged researchers from continuing to co-author papers with Russian scientists. The German Research Foundation (DFG) — the country's main funding agency — said that this didn't apply to work that had been completed or submitted to journals before the invasion, and a DFG spokesperson adds that the agency is not monitoring its recommendations or introducing punishments for researchers who go against them.



Ukrainian researchers pressure journals to boycott Russian authors

The research community has also discussed individual boycotts of Russian work — including in journals. International journals have generally not banned the consideration of Russian-authored work, and some researchers and journals have warned against indiscriminately isolating the country's scientists. But at least one title, Elsevier's *Journal of Molecular Science*, has said it will no longer consider manuscripts from scientists at Russian institutions. Its editor-in-chief, Rui Fausto, a chemist at the University of Coimbra in Portugal, says that the policy still stands, although *Nature*'s analysis found at least nine articles published in the journal that had Russian

authors and had been submitted after the war began. (Fausto says that he is looking into them.)

The Ukrainian government has strongly discouraged collaboration with Russian researchers and publication in Russian journals, says Michael Rose, who studies the economics of science and innovation at the Max Planck Institute for Innovation and Competition in Munich, Germany.

The war has come in a general period where nations have become more aware of the competitive geopolitical aspects of research collaboration, Flanagan adds, with countries expanding export controls and introducing restrictions on overseas collaboration in certain activities that have been deemed sensitive to national security.

Nature **615**, 199-200 (2023)

doi: https://doi.org/10.1038/d41586-023-00552-w

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00552-w

| Section menu | Main menu |

- NEWS FEATURE
- 08 March 2023
- Clarification 10 March 2023

In AI, is bigger always better?

As generative AI models grow larger and more powerful, some scientists advocate for leaner, more energy-efficient systems.

- Anil Ananthaswamy



Illustration by Fabio Buonocore

Artificial-intelligence systems that can churn out fluent text, such as OpenAI's ChatGPT, are the newest darlings of the technology industry. But when faced with mathematical queries that require reasoning to answer,

these large language models (LLMs) often stumble. Take, for instance, this algebra problem:

A line parallel to $y = 4x + 6$ passes through (5, 10). What is the y-coordinate of the point where this line crosses the y-axis?

Although LLMs can sometimes answer these types of question correctly, they more often get them wrong. In one early test of its reasoning abilities, ChatGPT scored just 26% when faced with a sample of questions from the 'MATH' data set of secondary-school-level mathematical problems[1].

This is to be expected: given input text, an LLM simply generates new text in accordance with statistical regularities in the words, symbols and sentences that make up the model's training data. It would be startling if just learning language patterns could allow LLMs to mimic mathematical reasoning reliably.

But back in June 2022, an LLM called Minerva, created by Google, had already defied these expectations — to some extent. Minerva scored 50% on questions in the MATH data set[2], a result that shocked some researchers in artificial intelligence (AI; see 'Minerva's mathematics test').

"The chatter in the community is that this is really kind of astounding," says Sébastien Bubeck, a machine-learning specialist at Microsoft Research in Redmond, Washington.

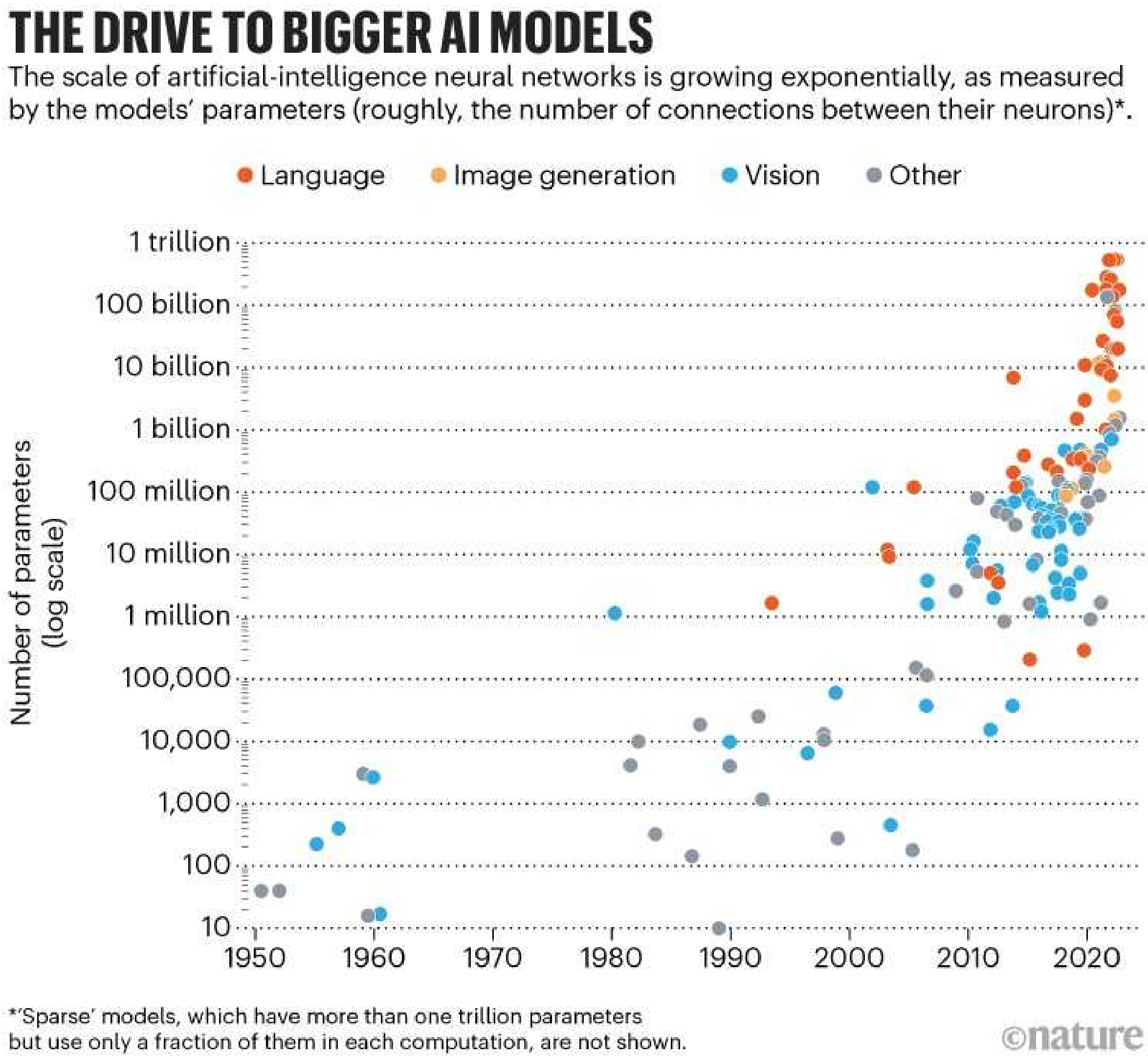
Minerva had the advantage that it was trained on mathematics-related texts. But Google's study suggested another important reason the model did so well — its sheer size. It was around three times the size of ChatGPT.

The Minerva results hint at something that some researchers have long suspected: that training larger LLMs, and feeding them more data, could give them the ability, through pattern-recognition alone, to solve tasks that are supposed to require reasoning. If so, some AI researchers say that this 'bigger is better' strategy might provide a path to powerful AI.

But there are reasons to doubt this thesis. LLMs still make glaring errors, and some scientists suggest that larger models are merely getting better at

answering queries that fall within the scope of the correlations in their training data, rather than acquiring an ability to answer entirely new questions.

The debate is now playing out on the frontiers of AI. Commercial firms have seen better results with bigger AI models, so they are rolling out ever-larger LLMs — each costing millions of dollars to train and run (see ‘The drive to bigger AI models’). But these models have major downsides. Besides concerns that their output cannot be trusted, and that they might exacerbate the spread of misinformation, they are expensive and suck up huge amounts of energy.



Source: Adapted from Our World in Data, and from J. Sevilla *et al*. Preprint at arXiv https://doi.org/10.48550/arXiv.2202.05924 (2022).

Critics argue that, ultimately, big LLMs will never be able to mimic or acquire skills that allow them to answer reasoning problems consistently. Instead, some scientists say, smaller, more energy-efficient AI is the way to make progress — inspired, in part, by the way the brain seems to learn and make connections.

Big, bigger, better?

LLMs such as ChatGPT and Minerva are giant networks of computing units (also called artificial neurons), arranged in layers. An LLM's size is measured in how many parameters it has — the adjustable values that describe the strength of the connections between neurons. Training such a network involves asking it to predict masked portions of a known sentence and tweaking these parameters so that the algorithm does a little better next time.



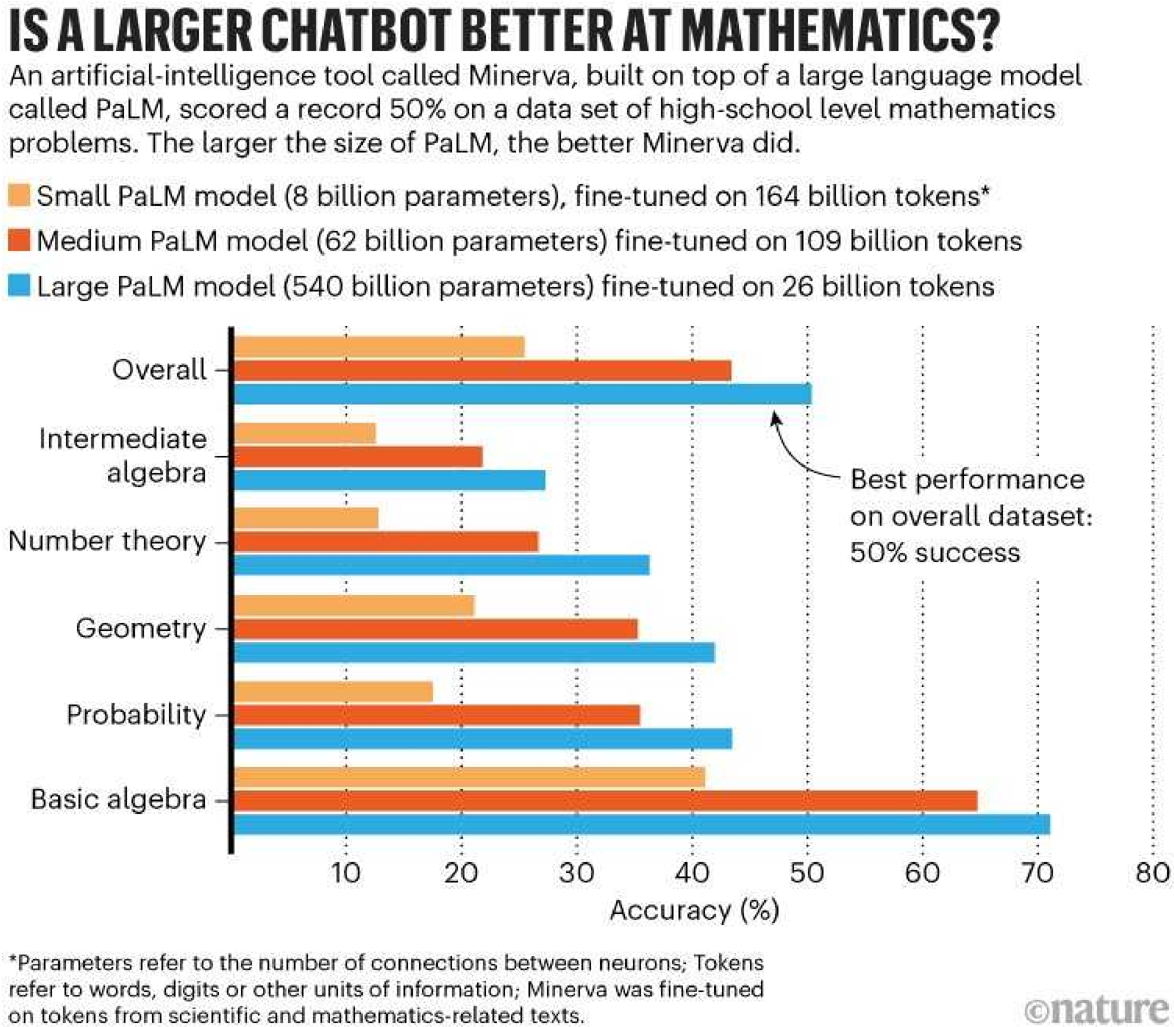
What ChatGPT and generative AI mean for science

Do this repeatedly over billions of human-written sentences, and the neural network learns internal representations that model how humans write language. At this stage, an LLM is said to be pre-trained: its parameters capture the statistical structure of written language that it saw during training, including all the facts, biases and errors in the texts. It can then be 'fine-tuned' on specialized data.

To make Minerva, for instance, researchers started with Google's Pathways Language Model (PaLM), which has 540 billion parameters and was pre-trained on a data set of 780 billion tokens[3]. A token can be a word, digit or some unit of information; in PaLM's case, the tokens were gleaned from English and multilingual web documents, books and code. Minerva was the result of fine-tuning PaLM on tens of billions of tokens from scientific papers and mathematics-related web pages.

Minerva can answer prompts such as: what is the largest multiple of 30 that is less than 520? The LLM appears to be thinking through the steps, and yet all it is doing is turning the questions into a sequence of tokens, generating a statistically plausible next token, appending it to the original sequence, generating another token, and so on: a process called inference.

Google researchers fine-tuned three sizes of Minerva using underlying pre-trained PaLM models of 8 billion, 62 billion and 540 billion parameters. Minerva's performance improved with scale. On the overall MATH data set, the smallest model had 25% accuracy, the medium-sized model reached 43% and the largest breached the 50% mark (see 'Is a larger chatbot better at mathematics?').



Source: Adapted from ref. 2

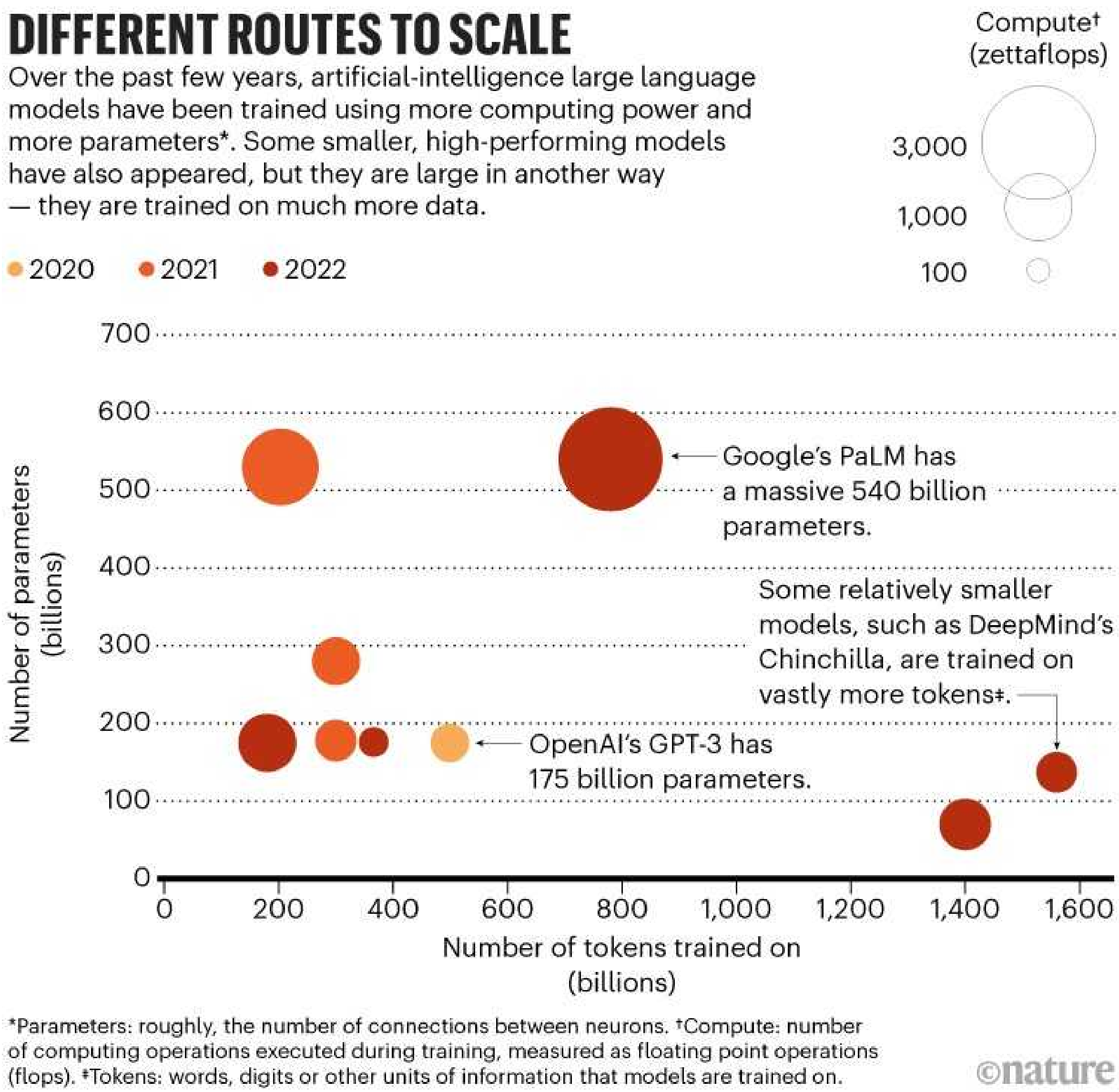
The biggest model also used the least amount of fine-tuning data — it was fine-tuned on only 26 billion tokens, whereas the smallest model looked at 164 billion tokens. But the biggest model took a month to fine-tune, on specialized hardware that had eight times as much computing capacity as used for the smallest model, which was fine-tuned for only two weeks. Ideally, the biggest model would have been fine-tuned on more tokens, says Ethan Dyer at Google Research in Mountain View, California, who is a member of the Minerva team; this could have led to an even better performance. But the team felt that the computational expense wasn't feasible.

Scaling laws

That the biggest Minerva model did best was in line with studies that have revealed scaling laws — rules that govern how performance improves with model size. A study in 2020 showed that models did better when given one of three things: more parameters, more training data or more 'compute' (the number of computing operations executed during training)[4]. Performance scaled according to a power law, meaning that it improved as some power of, for example, the number of parameters.

However, researchers don't exactly know why. "The laws are purely empirical," says Irina Rish, a computer scientist at the University of Montreal and Mila — Quebec Artificial Intelligence Institute in Montreal, Canada.

For the best results, the 2020 study suggested that as training data is doubled, model size should increase five times. Work last year slightly revised this. In March, the London-based AI firm DeepMind argued that it's best to scale up model size and training data together, and that smaller models trained on more data do better than bigger models trained on fewer data[5] (see 'Different routes to scale'). For example, DeepMind's Chinchilla model has 70 billion parameters, and was trained on 1.4 trillion tokens, whereas its 280-billion-parameter Gopher model, was trained on 300 billion tokens. Chinchilla outperforms Gopher on tasks designed to evaluate what the LLM has learnt.

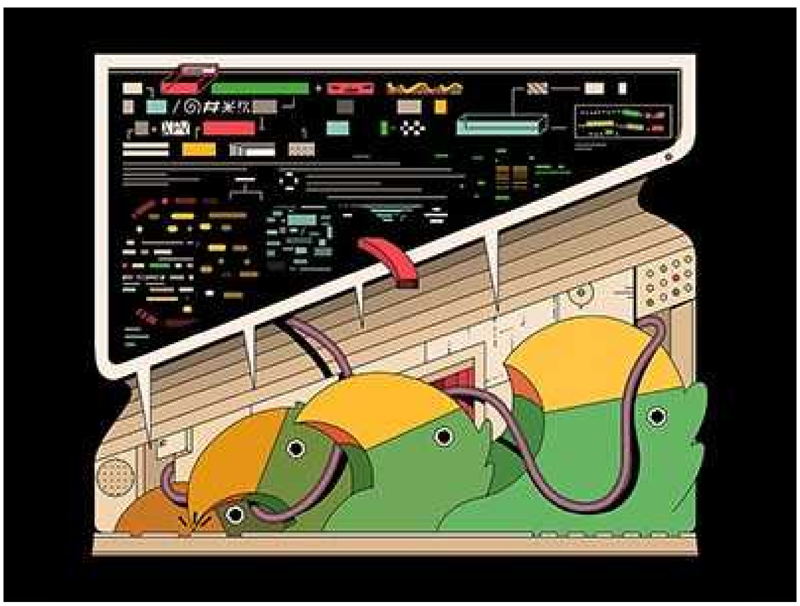


Source: Adapted from Our World in Data, and from J. Sevilla *et al*. Preprint at arXiv https://doi.org/10.48550/arXiv.2202.05924 (2022).

Scientists at Meta Research built on this concept in February with their own small-parameter model called LLaMA, trained on up to 1.4 trillion tokens. The 13-billion-parameter version of LLaMA outperformed ChatGPT's forerunner GPT-3 (175 billion parameters), the researchers say, whereas the 65-billion-parameter version was competitive with Chinchilla and even PaLM (see go.nature.com/3kje2fj).

Last October, Ethan Caballero at McGill University in Montreal, together with Rish and others, reported finding more complex relationships between size and performance[6]. In some instances, multiple power laws can govern how performance scales with model size, the researchers say.

For instance, in one hypothetical scenario fitting a general equation that they found, performance improves first gradually and then more rapidly with a model's size, but then dips slightly as the number of parameters continues to rise, before increasing again. The characteristics of this complex relationship are dictated by the specifics of each model and how it is trained. Ultimately, the researchers hope to be able to predict this in advance as any particular LLM is scaled up.



Robo-writers: the rise and risks of language-generating AI

A separate theoretical finding also supports the drive for bigger models — a 'law of robustness' for machine learning, introduced in 2021 by Bubeck and Mark Sellke, a statistician at Stanford University in California[7]. A model is robust if its answers remain consistent despite small perturbations in its inputs. Some AIs are notoriously fragile. If trained to recognize images of dogs, for instance, they will misclassify a test image when it's modified by a small amount of noise that wouldn't fool a person.

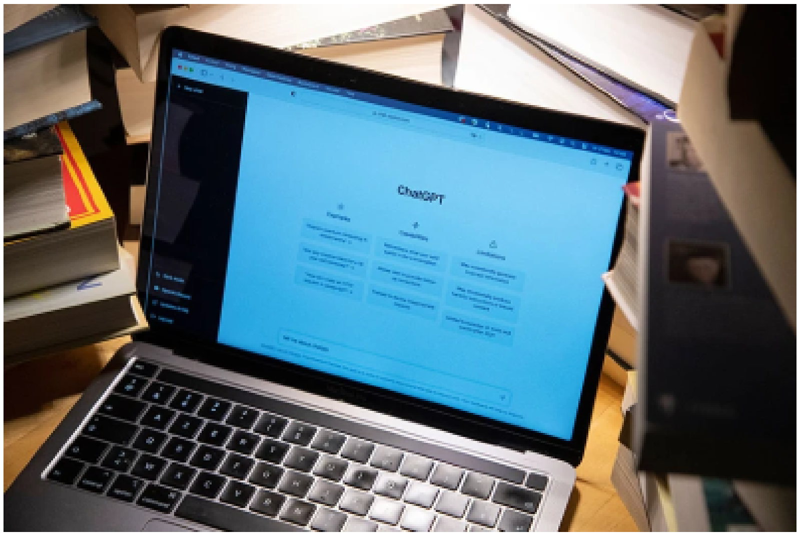
The more robust the AI, the better it generalizes to unseen data. Bubeck and Sellke have shown mathematically that increasing the number of parameters in a model increases robustness, and hence ability to generalize. The law proves that scaling up is necessary for generalization, but not that it's sufficient, says Bubeck. Nonetheless, it is being used to justify the move towards bigger models, he says. "Which I think is a reasonable thing."

Minerva also took advantage of a key innovation called chain-of-thought prompting. The user prefixes their question with text that includes a couple of examples of questions and solutions, including the reasoning — illustrating a typical chain of thought — that led to the answers. During inference, the LLM takes its cues from this context and provides a step-by-step answer that looks surprisingly like reasoning. This doesn't require updates to the model's parameters, and so doesn't involve the additional computing power that fine-tuning needs.

The ability to respond to chain-of-thought prompts shows up only in LLMs with more than about 100 billion parameters. Such discoveries have helped bigger models to improve in accordance with empirical scaling laws, says Blaise Agüera y Arcas at Google Research in Seattle, Washington. "The bigger models keep doing better and better."

Reasonable concerns

François Chollet, an AI researcher at Google in Mountain View, is among the sceptics who argue that no matter how big LLMs become, they will never get near to having the ability to reason (or mimic reasoning) well enough to solve new problems reliably. An LLM only appears to reason by using templates that it has encountered before, whether in the training data or in the prompt, he says. "It cannot, on the fly, make sense of something that it has not seen before."



[How Nature readers are using ChatGPT]

The best that LLMs might be able to do is to slurp in so much training data that the statistical patterns of language alone allow them to respond to questions with answers that are very close to what they've already seen.

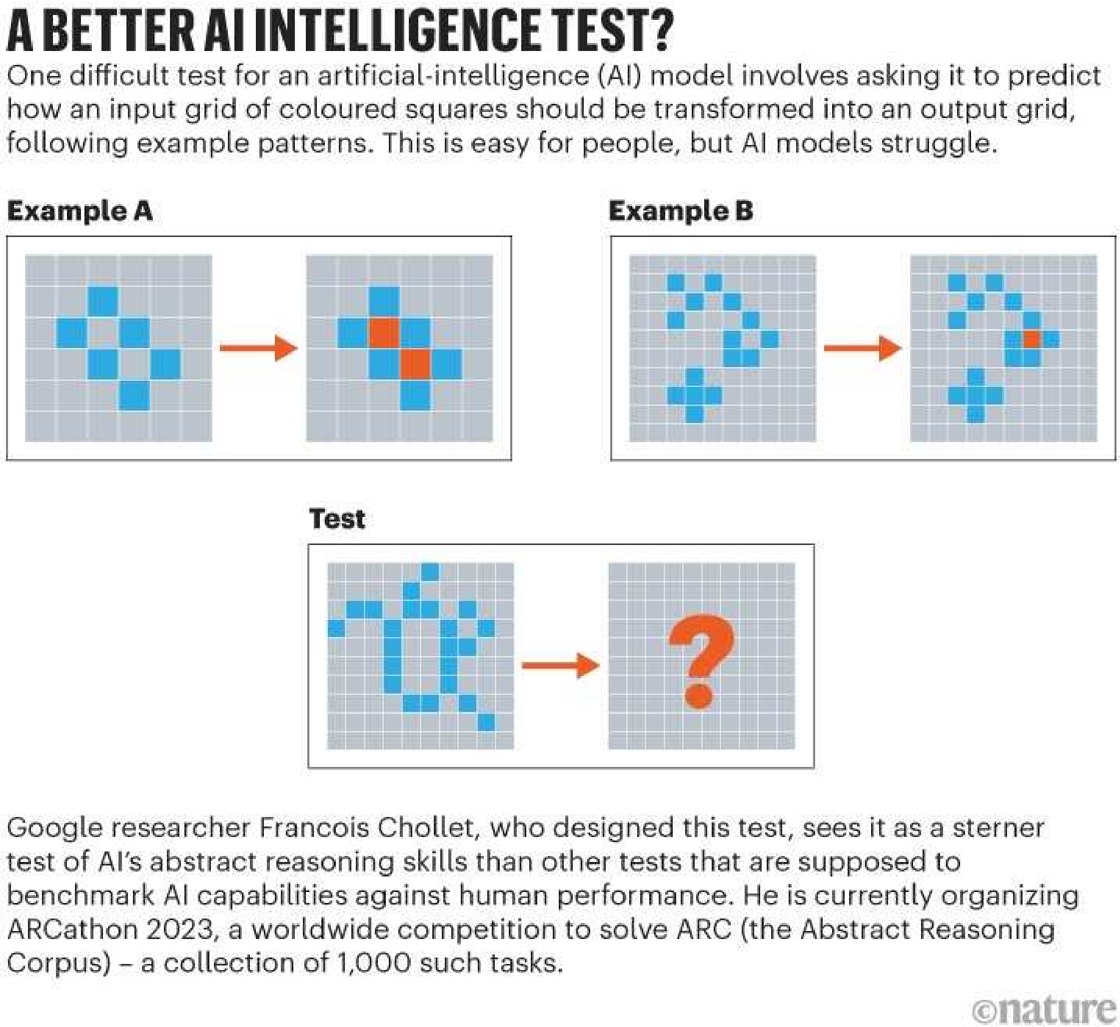
Agüera y Arcas, however, argues that LLMs do seem to have gained some surprising abilities that they weren't trained on specifically. In particular, he points to tests designed to show whether a person has what's called theory of mind — being able to theorize or gauge the mental states of others. Take this simple example. Alice puts her glasses away in a drawer. Then Bob, unbeknown to Alice, hides the glasses under a cushion. Where will Alice look for her glasses first? A child asked this question is being tested on whether they understand that Alice has her own beliefs that might not agree with what the child knows.

In his experiments with another of Google's LLMs called Language Model for Dialogue Applications (LaMDA), Agüera y Arcas found that LaMDA responded correctly in more extended conversations of this type. To him, this was suggestive of an LLM's ability to internally model the intentions of others. "These models that are doing nothing but predicting sequences develop an extraordinary range of capabilities, including theory of mind," says Agüera y Arcas (see 'Theory of mind?'). But he concedes that these models are error-prone, and he, too, is unsure whether scaling alone, although it seems necessary, is sufficient for reliable reasoning.

Even when LLMs get the answers right, however, there is no understanding involved, says Chollet. “When you try to probe it a little bit, it becomes immediately obvious that it’s all empty. ChatGPT has no model of what it is talking about,” he says. “You’re watching a puppet show and believing the puppets are alive.”

So far, LLMs still make absurd mistakes that humans never would, says Melanie Mitchell, who studies conceptual abstraction and analogy-making in AI systems at the Santa Fe Institute in New Mexico. This has contributed to the many concerns about the safety of unleashing LLMs without guardrails into society.

An issue with the debate over whether LLMs can ever tackle genuinely new, unseen problems is that we have no way of comprehensively testing for this ability, Mitchell adds. “The current benchmarks we have are not adequate,” she says. “They’re not probing things systematically. We don’t really know yet how to do that.” Chollet advocates for a test of abstract reasoning that he designed, called the Abstract Reasoning Corpus (see ‘A better AI intelligence test?’).



Source: F. Chollet/Lab42 (https://arc-editor.lab42.global)

The problems of scale

While the debate plays out, there are already pressing concerns over the trend towards larger language models. One is that the data sets, computing power and expense involved in training big LLMs restricts their development — and therefore their research direction — to companies with immense computing resources. OpenAI has not confirmed the costs of creating ChatGPT, but others have estimated on the basis of the compute involved that pre-training GPT-3 (a predecessor of ChatGPT) would have cost more than US$4 million. It's probably costing OpenAI millions of

dollars each month to run ChatGPT, because of the number of queries that the free chatbot is now servicing. “We are already deep into this regime,” says Bubeck. “There are just a few companies that have models bigger than 100 billion parameters.”

Governments are starting to step in with support that might expand the playing field. In June last year, an international team of around 1,000 academic volunteers, with funding from the French government, a US AI company called Hugging Face and others, trained a model called BLOOM with about 175 billion parameters, using $7 million worth of computing time[8]. And in November, the US Department of Energy awarded supercomputing time to a project from Rish and her colleagues, to build large models to study their behaviour. “We hope to train a Chinchilla-like 70-billion-parameter model — not necessarily the largest, but rather the one whose performance scales more effectively,” says Rish.

Regardless of who gets to build them, LLMs also raise concerns about electricity consumption. For example, Google reported that training PaLM took about 3.4 gigawatt-hours over about two months. That’s the annual energy consumption of about 300 US households. Google trained PaLM at its Oklahoma data centre, which it said operated on 89% carbon-free energy, being powered mostly by wind and other renewable sources. But a survey of industry AI models has shown that the majority are trained using electricity grids that are still largely powered by fossil fuels[9].

Chollet’s concern is that as multiple firms begin training and using bigger models, they could start to suck up more electricity. “Every big tech company will now attempt to deploy LLMs into their products, regardless of whether it’s a good idea or not,” he says.

Smarter and smaller?

For many scientists, then, there’s a pressing need to reduce LLM’s energy consumption — to make neural networks smaller and more efficient, as well as, perhaps, smarter. Besides the energy costs of training LLMs (which, although substantial, are a one-off), the energy needed for inference — in which LLMs answer queries — can shoot up as the number of users

increases. Big tech firms haven't commented on the usage costs of their models. Hugging Face, however, revealed that when its BLOOM model was deployed on the Google Cloud Platform for 18 days, in which time it answered 230,768 queries (many fewer than ChatGPT, which hit 100 million active users a month in February), it consumed, on average, 1,664 watts[10].

For comparison, our own brains are much more complicated and larger than any LLM, with 86 billion neurons and some 100 trillion synaptic connections. And yet, the human brain consumes somewhere between 20 and 50 watts of power, says Friedemann Zenke at the Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland.



Open-source language AI challenges big tech's models

So some researchers hope that mimicking aspects of the brain will help LLMs and other neural networks to become smaller, smarter and more efficient.

One source of the brain's overall intelligence and efficiency might be its recurrent, or feedback, connections. LLMs are, essentially, 'feedforward' networks. This means that information flows one way: from the input, through the layers of the LLM, to its output. The brain is wired differently. For example, in the human visual system, neurons connect regions of the brain that first receive visual information to areas further back. But there are

also feedback connections that allow information transfer between neurons in the reverse direction. “There’s maybe ten times as many feedback connections as there are feedforward connections in the [human] visual system,” says Mitchell, but an LLM has no feedback connections.

Artificial neural networks that incorporate both feedforward and feedback connections are generically called recurrent neural networks (RNNs). Such networks (unlike feedforward LLMs) can discern patterns in data that change over time. That’s “fundamental to how all natural intelligences experience the world and learn”, says Kanaka Rajan, a computational neuroscientist at the Icahn School of Medicine at Mount Sinai in New York City. But RNNs come with challenges, says Rajan. For instance, they are hard and slow to train, making it difficult to scale them to the size of current LLMs.

Another reason brains are efficient is that biological neurons mostly remain quiet — they have only an occasional spike in activity. By contrast, the artificial neurons in most neural networks are modelled as being constantly on. Researchers are studying artificial neurons that spike (mimicking real ones), but it is difficult to adapt algorithms that train standard neural networks to networks that use spiking neurons. Still, research using small data sets (for example, 10,000 audio recordings used to train a network to recognize spoken digits) has shown that RNNs with spiking neurons outperform those with standard neurons, and, in theory, are three orders of magnitude more computationally efficient[11]. “Progress is rapid and impressive,” says Sander Bohté, at Amsterdam’s National Research Institute for Mathematics and Computer Science in the Netherlands (CWI), who works in this area.

As long as such spiking networks are only simulated in software, however, they cannot provide real efficiency gains (since the hardware simulating them still consumes power). Such computing elements will need to be built into hardware, on neuromorphic chips, to realize their benefits.

Energy-efficient LLMs

Meanwhile, researchers are experimenting with different ways to make existing LLMs more energy efficient, and smarter. In December 2021, DeepMind reported a system called RETRO, which combines an LLM with an external database. The LLM uses relevant text retrieved from this database during inference to help it make predictions. DeepMind's researchers showed that a 7.5-billion parameter LLM, coupled with a database of 2 trillion tokens, outperforms LLMs with 25 times more parameters[12]. The researchers wrote that this was a "more efficient approach than raw parameter scaling as we seek to build more powerful language models".

In the same month, scientists at Google Research reported another way to increase energy efficiency at scale. Their Generalist Language Model, or GLaM, has 1.2 trillion parameters[13]. But these parameters don't represent one giant neural network; internally, they are distributed between 64 smaller neural networks, alongside other layers. The LLM is trained such that during inference, it uses only two of its networks to complete a task; overall, the LLM uses only about 8% of its trillion-plus total parameters for inference, per token. According to Google, GLaM used the same amount of computing resources as were needed to train GPT-3, but consumed only about one-third of the power, because of improvements in training software and hardware. During inference, GLaM used half the computing resources that GPT-3 needed. And it outperformed GPT-3 when trained on the same amount of data.

To improve further, however, even these more energy-efficient LLMs seem destined to become bigger, using up more data and compute. Researchers will be watching to see what new behaviours emerge with scale. "Whether it will fully unlock reasoning, I'm not sure," says Bubeck. "Nobody knows."

Nature **615**, 202-205 (2023)

doi: https://doi.org/10.1038/d41586-023-00641-w

Updates & Corrections

- **Clarification 10 March 2023**: This article has been modified to clarify that Mark Sellke, a statistician at Stanford University in California, was

the only other person involved in introducing the robustness law.

References

1. Frieder, S. *et al.* Preprint at https://arxiv.org/abs/2301.13867 (2023).
2. Lewkowycz, A. *et al.* Preprint at https://arxiv.org/abs/2206.14858 (2022).
3. Chowdhery, A. *et al.* Preprint at https://arxiv.org/abs/2204.02311 (2022).
4. Kaplan, J. *et al.* Preprint at https://arxiv.org/abs/2001.08361 (2020).
5. Hoffmann, J. *et al.* Preprint at https://arxiv.org/abs/2203.15556 (2022).
6. Caballero, E. *et al.* Preprint at https://arxiv.org/abs/2210.14891 (2022).
7. Bubeck, S. *et al.* Preprint at https://arxiv.org/abs/2105.12806 (2021).
8. Le Scao, T. *et al.* Preprint at https://arxiv.org/abs/2211.05100 (2022).
9. Luccioni, A. S. & Hernandez-Garcia, A. Preprint at https://arxiv.org/abs/2302.08476 (2023).
10. Luccioni, A. S., Viguier, S. & Ligozat, A.-L. Preprint at https://arxiv.org/abs/2211.02001 (2022).
11. Yin, B. *et al. Nature Mach. Intell.* **3**, 905–913 (2021).
12. Borgeaud, S. *et al.* Preprint at https://arxiv.org/abs/2112.04426 (2021).
13. Du, N. *et al.* Preprint at https://arxiv.org/abs/2112.06905 (2021).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00641-w

- NEWS FEATURE
- 07 March 2023

Indoor air is full of flu and COVID viruses. Will countries clean it up?

The current pandemic has focused attention to the importance of healthy indoor air and could spur lasting improvements to the air we breathe.

- Dyani Lewis



Dance clubs and other indoor spaces in Belgium will soon post information about air quality.Credit: Christophe Ketels/Alamy

Bars in Belgium could be among the healthiest places to have a drink, come July. That's when a new law goes into effect, requiring public venues to meet air-quality targets and display real-time measurements of carbon dioxide concentrations — a proxy for how much clean air is piped in.

Consumers in Belgium will get even more information in 2025, when gyms, restaurants and indoor workspaces must all show air-quality ratings given through a certification system. In the event of a future pandemic, Belgium's rating system could determine whether or not a venue is closed.

The law, enacted in July 2022, is the boldest in a string of moves that countries have taken in the wake of the COVID-19 pandemic to make indoor spaces safer in the face of infectious diseases caused by viruses such as SARS-CoV-2 and influenza.

In March 2022, the US government launched a Clean Air in Buildings Challenge to spur building owners and operators to improve their ventilation and indoor air quality. In October last year, the state of California passed a law requiring all school buildings to provide clean indoor air. And in December, the White House announced that all federal buildings — some 1,500 in total — would meet minimum air-safety requirements. Also in December, the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) — a construction-industry body whose recommendations are adopted into law through local building codes in the United States and elsewhere — announced that it would be developing standards that take infection risk into account by June 2023.



Hidden harms of indoor air pollution — five steps to expose them

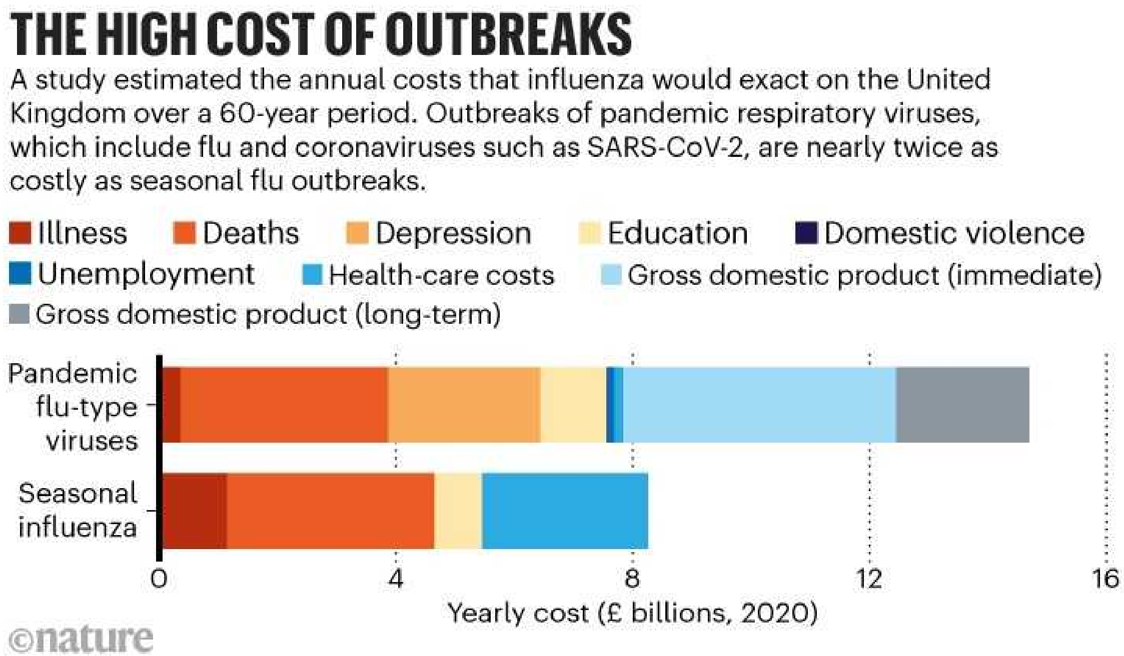
Last June, the United Kingdom's leading engineering bodies released a report, commissioned by the government, that called for enforceable clean-air regulations to make buildings safe over their entire lifetimes (see go.nature.com/3kgsmjt). Other countries are also taking steps — for example, by deploying air-quality monitors in classrooms.

Specialists in indoor air quality are buoyed by the prospect that the pandemic could bring lasting improvements to the air we breathe indoors. The SARS-CoV-2 virus that causes COVID-19 is spread mainly in indoor spaces, as are the pathogens that lead to other infectious diseases, such as chicken pox, measles, tuberculosis and seasonal influenza.

"There's never been, in history, so much action about indoor air quality," says Lidia Morawska, an aerosol scientist at the Queensland University of Technology in Brisbane, Australia.

But huge challenges lie ahead, particularly for the existing stock of schools, office buildings and other public venues. Retrofitting them with the technology to deliver clean air at sufficient levels will be an immense — and costly — undertaking, say experts in this field. But, they argue, the benefits would outweigh the costs. By one estimate, pandemic and seasonal influenza outbreaks cost the United Kingdom £23 billion (US$27 billion) per year, on

average (see 'The high cost of outbreaks'), and the country could save £174 billion over a 60-year period by improving ventilation in buildings (see go.nature.com/3ktumeg).



Source: NERA Economic Consulting

Making indoor spaces safe from infection could also reduce exposure to pollutants such as fine particulates from wildfire smoke and cooking, volatile organic compounds leached from furniture, and allergy-causing moulds and pollen. But it could also raise energy costs and contribute to greenhouse-gas emissions.

Researchers are still working to pin down how best to ventilate indoor spaces to prevent infections from spreading, and what alternative technologies might replace or enhance mechanical ventilation systems. But many say that enough is already known to start demanding safer indoor spaces.

It's a race against time. As concern over COVID-19 wanes, experts wonder how much progress countries will make before the next big outbreak of an airborne infectious disease.

Reducing infections

When COVID-19 reached pandemic status in early 2020, health officials didn't pay much attention to the risks of indoor air. Initially, the World Health Organization (WHO) dismissed the role of airborne transmission and focused — incorrectly — on transmission through contaminated surfaces. But even when public-health authorities began recommending better ventilation as a way of preventing infection, they offered only vague guidance. Authorities told people to open windows and bring in as much outdoor air as possible with mechanical ventilation systems, without giving specific numbers.

Such advice sowed confusion, says Joseph Allen, a building hygienist at the Harvard T.H. Chan School of Public Health in Boston, Massachusetts. "You can't tell people to bring in more outdoor air without answering how much," he says.



Why the WHO took two years to say COVID is airborne

Allen was one of the first to put a value on how much ventilation people should be aiming for. In June 2020, he and his colleagues recommended that schools wanting to reopen their doors after lockdowns should deliver four to six air changes per hour to their classrooms[1] — changes in which the entire

volume of air in the room is replaced. That amounts to a ventilation rate of 10–14 litres per second per person. Most schools were achieving much less than that, however. A study of California classrooms, for example, found that most failed to meet that level of ventilation[2]. The WHO issued its own guidelines in March 2021, recommending a ventilation rate of 10 litres per second per person outside health-care settings.

In theory, the pandemic provided the perfect opportunity to gather real-world data to see whether low ventilation rates were associated with outbreaks, and to test different rates of ventilation to see which resulted in reduced infection rates. But health officials only rarely considered ventilation when investigating major outbreaks of COVID-19. Yuguo Li, a mechanical engineer at the University of Hong Kong, estimates that fewer than ten investigations measured ventilation rates in venues where outbreaks occurred, because airborne transmission was not on people's radar.

Instead, researchers tried to gain clues through observational studies. Morawska was involved in one that looked at 10,000 school classrooms in the Marche region of Italy. In the 316 classrooms that had mechanical ventilation with rates of 1.4–14 litres per second per person, the students' risk of infection was reduced by at least 74% over a 4-month period at the end of 2021, compared with that for students in classrooms that relied on windows for ventilation. This group typically received less than 1 litre per second per person. When ventilation rates were at least 10 litres per second per student, the infection risk was 80% lower[3].



Why indoor spaces are still prime COVID hotspots

Evidence is also growing about other technologies that remove infectious particles from the air. One study[4] explored the effectiveness of two air cleaners fitted with high-efficiency particulate absorbing (HEPA) filters, placed in a 54-square-metre conference room with a dummy that generated aerosol particles similar to those that transmit SARS-CoV-2. The cleaners reduced the aerosol exposure of three dummy participants by 65%. That's just shy of the 72% reduction achieved by masking all of the dummy participants[4].

Another study, by civil engineer Bert Blocken at the Catholic University of Leuven (KU Leuven) in Belgium, found that ventilation combined with air cleaning, equivalent to 6 air changes per hour in total, reduced exhaled aerosol concentrations in a gym to 5–10% of what they would have been without these measures[5]. That concentration substantially reduces infection risk, says Blocken. He adds that air cleaners are an underappreciated technology that could be readily deployed in buildings that don't have mechanical ventilation systems capable of providing enough clean air, or where operating such systems would consume too much energy. The state of Victoria in Australia took this approach, distributing portable air cleaners to all of its 110,000 classrooms in 2022.

Last November, the *Lancet* COVID-19 Commission's Task Force on Safe Work, Safe School, and Safe Travel, chaired by Allen, published concrete guidelines for clean-air delivery rates — using ventilation, air filtration or other means — to reduce airborne infections[6]. To achieve what the report describes as the 'best' air quality, it recommends more than 6 air changes per hour, or 14 litres per second per person (see 'How much clean air is enough?').

How much clean air is enough?

A task force of researchers proposed ventilation rates for buildings using several metrics*, with the aim of reducing the risks of transmission of airborne respiratory diseases.

Assessment	Equivalent air exchanges per hour	Cubic feet per minute per person	Litres per second per person
Good	4	21	10
Better	6	30	14
Best	>6	>30	>14

*Not shown: volumetric flow rates per floor area.

Legal limits

Ventilation requirements can be complicated, because they change depending on how big the space is, how many people are in it and how active they are. So some researchers advocate using a shortcut — setting maximum carbon dioxide concentrations. CO_2 is frequently used as a proxy measure for ventilation and indoor air quality[7]. Because people exhale CO_2 as they breathe, levels of the gas can shoot up if a space is crowded or if there is insufficient ventilation to replace the exhaled air — which might contain infectious viruses — with clean air.

Until 1999, ASHRAE standards included a recommended limit for CO_2 of 1,000 parts per million (p.p.m.). At this concentration, according to research conducted in the 1930s, building occupants' perception of body odour would be kept at an acceptable level. Since then, research has shown that when concentrations exceed 1,000 p.p.m., CO_2 can cause drowsiness and can impair cognitive performance on decision-making and problem-solving tasks[8].

A small study published in September 2022 — and yet to be peer reviewed — directly connected CO_2 levels with those of infectious pathogens. The

authors tested air samples in nurseries, schools, universities and care homes for the presence of respiratory pathogens. Rooms that had higher CO_2 levels were associated with higher levels of respiratory pathogens[9].

In August 2021, the UK government began distributing CO_2 sensors to all school classrooms so that teachers could use the devices to decide when to open windows or increase ventilation. Similar schemes have been rolled out in Europe, the United States and elsewhere, although none has yet been evaluated for its ability to reduce infection rates.

Relying on CO_2 readings has drawbacks, however. Concentrations can creep up even when the infection risk remains low, such as when using portable air cleaners — which do not remove CO_2 from the air — or when cooking. CO_2 is useful, says chemist Nicola Carslaw at the University of York, UK, who studies indoor-air pollutants, “but it’s definitely not the whole story”.



A green lamp indicator shows indoor carbon dioxide levels are safe in a lecture hall at the University of Duisberg-Essen, Germany.Credit: dpa picture alliance/Alamy

Despite these issues, Morawska says that CO_2 monitors should be widely deployed as an inexpensive, readily available tool that could be installed in every indoor space, much like smoke alarms. But displaying CO_2 read-outs on its own is not enough, she adds, because it places the onus on room occupants to track air quality and decide what to do if readings are high.

Morawska would also like to see laws that mandate maximum CO_2 levels permissible in public buildings, so that the responsibility is placed back on building operators and government regulators. A handful of governments have already done just that. Last year, Morawska and her colleague Wei Huang at Peking University in Beijing reviewed air-quality laws in more than 100 countries. Only 12 had national standards for indoor air quality that specified threshold limits for pollutants. And only 8 of those — including China, South Korea, India, Poland and Hungary — set limits for CO_2 concentration, most between 800 p.p.m. and 1,000 p.p.m.[10].

Japan has had a law to regulate indoor air quality since 1970, which mandates that buildings must not exceed indoor CO_2 concentrations of 1,000 p.p.m.. The law requires that building managers assess air quality every two months, report results to the government and establish remediation plans if the air quality does not meet the standards. But almost 30% of buildings exceeded the CO_2 limit in 2017, according to a 2020 report[11].



[How record wildfires are harming human health]

Still, the Japanese laws work, says Kazukiyo Kumagai, a public-health engineer at the California Department of Health in Richmond. “Japan is in a better condition” than the United States when it comes to indoor air quality, he says. Adopting a Japanese-style approach of regular monitoring and reporting might work elsewhere, he adds.

Legal limits could become more common. The new Belgian law, for example, comes into effect in July this year and stipulates that public venues ventilate at a rate of 40 cubic metres per hour so that CO_2 does not exceed 900 p.p.m.. If air filtration is used, a lower ventilation rate of 25 cubic metres per hour is enough, and CO_2 can reach a maximum level of 1,200 p.p.m..

Legislating indoor air quality is “tricky” says Catherine Noakes, a mechanical engineer at the University of Leeds, UK, who contributed to that country’s report into infection-resilient buildings. “One of the challenges with indoor air,” she says, “is who owns it?” The responsibility can be distributed across government departments and agencies, depending on how the building is used. A school’s indoor air might be the responsibility of the education department, whereas office buildings could be regulated by an occupational health and safety agency.

That's the situation in the United States, where no agency currently has the authority to regulate indoor air, says Andrew Persily, a mechanical engineer at the National Institute of Standards and Technology in Gaithersburg, Maryland. In Belgium, too, the new national law doesn't cover schools, which are the responsibility of regional governments. And in Japan, a separate law for school buildings specifies a higher CO_2 limit of 1,500 p.p.m., a level many regard as too high.

Setting standards

In the absence of national laws, professional bodies that set air-quality standards are starting to act. When ASHRAE releases its infection-mitigation standard in June, the hope is that these recommended targets will be adopted into local building codes that new buildings must comply with.

"We have always addressed indoor air quality, but not specifically for pathogen mitigation," says engineer Ginger Scoggins, the president-elect of ASHRAE, who is based in North Carolina. ASHRAE could face some pushback. Scoggins says that when the society made a previous change to increase the ventilation requirement from 5 cubic feet per minute to 15 (2.4 litres per second to 7.1 litres per second), many people in the warm parts of the United States were angry because it would drive up energy costs from air conditioning. Her local school board passed a ruling that its classrooms only needed to get to 7.5.

Even though ASHRAE standards are not enforced, they will make a difference, says Allen. Aside from influencing how buildings are constructed, more stringent ASHRAE standards send a strong signal to businesses in older buildings about what the gold standard for indoor air quality looks like.

An economic case could be made for better indoor air, says Noakes. The cost–benefit analysis conducted for the UK report found that the country could save £3 billion per year over a 60-year period by improving ventilation.

Researchers say it will take time to lower the infection risks inside buildings. “We are looking at 30 years,” says Morawska. “But we are talking about the future of our society.”

Nature **615**, 206-208 (2023)

doi: https://doi.org/10.1038/d41586-023-00642-9

References

1. Jones, E. *et al. Schools for Health: Risk Reduction Strategies for Reopening Schools* (Harvard T.H. Chan School of Public Health Healthy Buildings Program, 2020).

2. Mendell, M. J. *et al. Indoor Air* **23**, 515–528 (2013).

3. Buonanno, G. *et al. Front. Public Health* **10**, 1087087 (2022).

4. Lindsley, W. G. *et al. Morb. Mortal. Wkly Rep.* **70**, 972–976 (2021).

5. Blocken, B. *et al. Build. Environ.* **193**, 107659 (2021).

6. The Lancet COVID-19 Commission Task Force on Safe Work, Safe School, and Safe Travel. *Proposed Non-infectious Air Delivery Rates (NADR) for Reducing Exposure to Airborne Respiratory Infectious Diseases* (Lancet COVID-19 Commission, 2022).

7. Peng, Z. & Jimenez, J. L. *Environ. Sci. Technol. Lett.* **8**, 392–397 (2021).

8. Wargocki, P., Porras-Salazar, J. A., Contreras-Espinoza, S. & Bahnfleth, W. *Build. Environ.* **173**, 106749 (2020).

9. Raymenants, J. *et al.* Preprint at medRxiv https://doi.org/10.1101/2022.09.23.22280263 (2022).

10. Morawska, L. & Huang, W. In *Handbook of Indoor Air Quality* (eds Zhang, Y. *et al.*) (Springer, 2022).

11. Hayashi, M., Kobayashi, K., Kim, H. & Kaihara, N. *J. Natl Inst. Public Health* **69**, 63–72 (2020).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00642-9

Books & Arts

- **Seven everyday objects that made the modern world** [06 Mar 2023]
  • Nails, wheels, springs, magnets, lenses, string and pumps: a structural engineer reveals the small things that our biggest tech advances are built on.
- **The history of profit, and are animals creative? Books in brief** [27 Jan 2023]
  • Andrew Robinson reviews five of the best science picks.

- BOOK REVIEW
- 06 March 2023

Seven everyday objects that made the modern world

Nails, wheels, springs, magnets, lenses, string and pumps: a structural engineer reveals the small things that our biggest tech advances are built on.

- Anna Novitzky



Centuries of development led to the bicycle's spoked wheel — which brought mobility and freedom for many.Credit: Getty

Nuts and Bolts: Seven Small Inventions That Changed the World (in a Big Way) *Roma Agrawal* Hodder & Stoughton (2023)

Ten thousand years ago, the most advanced tools were made of chipped stone. Today, much technology is so complicated that it's almost "indistinguishable from magic", to borrow Arthur C. Clarke's phrase. How did we get here? In *Nuts and Bolts*, structural engineer Roma Agrawal investigates the story of basic technological developments, and shows how intimately entwined they are with humanity's own history.

She contends that the modern world has its foundations in seven humble inventions: the nail, wheel, spring, magnet, lens, string and pump. Building on her lifelong fascination with opening things up to see what makes them work, she explores the science of each of these mini 'machines', and follows their history from ancient beginnings to manifestations of modern engineering, small and large.

The spring, for example, she characterizes as "humanity's first tool that allowed us to store energy and then release it when we wanted". She charts its development from bows and arrows to the vast steel coils that help skyscrapers to withstand earthquakes, and the silicon hairsprings that maintain the accuracy of the most exclusive mechanical watches.



The bicycle problem that nearly broke mathematics

With a clear, lively and engaging style — and many puns — Agrawal encourages a new perspective on the inventions that keep the world rolling. The wheel, she explains, was first used not for transport, but to make pottery. It wasn't for another 700 years that it was turned on its side and attached to an axle; the earliest surviving wheeled vehicles are from around 3200 bc, in what's now Russia. A multitude of refinements followed, technological advances ushering in sweeping social change. Spoked wheels supplanted solid ones, their lightness enabling fast travel and improved trade. Wheels with wire spokes led to the bicycle, a source of freedom for many who couldn't afford carriages or cars.

Gear change

Serrated wheels became gears, tools for changing the direction and magnitude of forces that prompted advances of the Industrial Revolution and beyond. In the late nineteenth century, for instance, Josephine Cochran used wheels and gears to develop the first commercially viable dishwasher — one of myriad appliances that would cut the time spent on housework and free many women to enter the workforce. And wheels are ancestors of the gyroscopes that steer the International Space Station. As Agrawal says, 5,000 years of change have been driven by constant reinvention of the wheel.

Refreshingly, this world-trotting tale focuses on "the often hidden or unacknowledged contribution of minoritised people". Soviet women work as fighter pilots in the Second World War; spinning wheels take a key role in the Indian independence movement. Even familiar names have an unexpected spin. Telephone inventor Alexander Graham Bell, for example, originally set out to convert the vibrations of speech into something that his mother and wife, both deaf, could see.



Degrowth can work — here’s how science can help

The global tapestry is shot through with Agrawal’s personal stories. She makes her own nails at a blacksmith’s forge, and describes movingly how she owes the birth of her daughter to the lens, without which the microscopes that were crucial to the development of *in vitro* fertilization (IVF) would not have existed. Her family history illustrates how using magnets to modulate electric pulses enabled technologies that have repeatedly revolutionized global communications. Her uncle, looking for work in Italy in the 1960s, sends telegrams to his family in Mumbai, India. An aunt in Connecticut in the 1970s books a trunk call across continents through a network of telephone exchanges. Agrawal’s daughter speaks her early words to her own grandmother by video call during the COVID-19 pandemic.

All that glitters

The reader can’t help but be swept along in Agrawal’s enthusiasm, but she’s at pains to point out that ‘progress’ isn’t always good. Spoked wheels made chariots nimble enough to be used in war. Springs are crucial to guns. String led to a textile industry that bankrolled the British Empire while laying waste to the Indian economy (hence the significance of the spinning wheel).

The fashion industry now generates around 10% of global carbon dioxide emissions and vast amounts of other pollution.

The question of engineering's impact on the planet and on society is a common thread. By drawing out stories of technology's use for good and ill, Agrawal encourages deep thinking about where we go from here.

She shows how, to manage humanity's devastation of the planet, we must demystify the 'black box' of technology. "Understanding the nuts and bolts of our objects", she says, leads us to appreciate the effort, ingenuity and raw materials that went into them, and stop seeing them as disposable. We become more likely to find ways to extend their lives, to reconstitute and repair them. Agrawal points to sustainable engineering practices, community-awareness programmes and repair movements as ways to help fend off environmental disaster. On an individual level, taking ownership of our possessions through repair brings a fulfilling "sense of happiness, satisfaction, and achievement".

So the history of engineering tells us about who we are and who we have been. It can even lead us to who we want to be.

Nature **615**, 209-210 (2023)

doi: https://doi.org/10.1038/d41586-023-00643-8

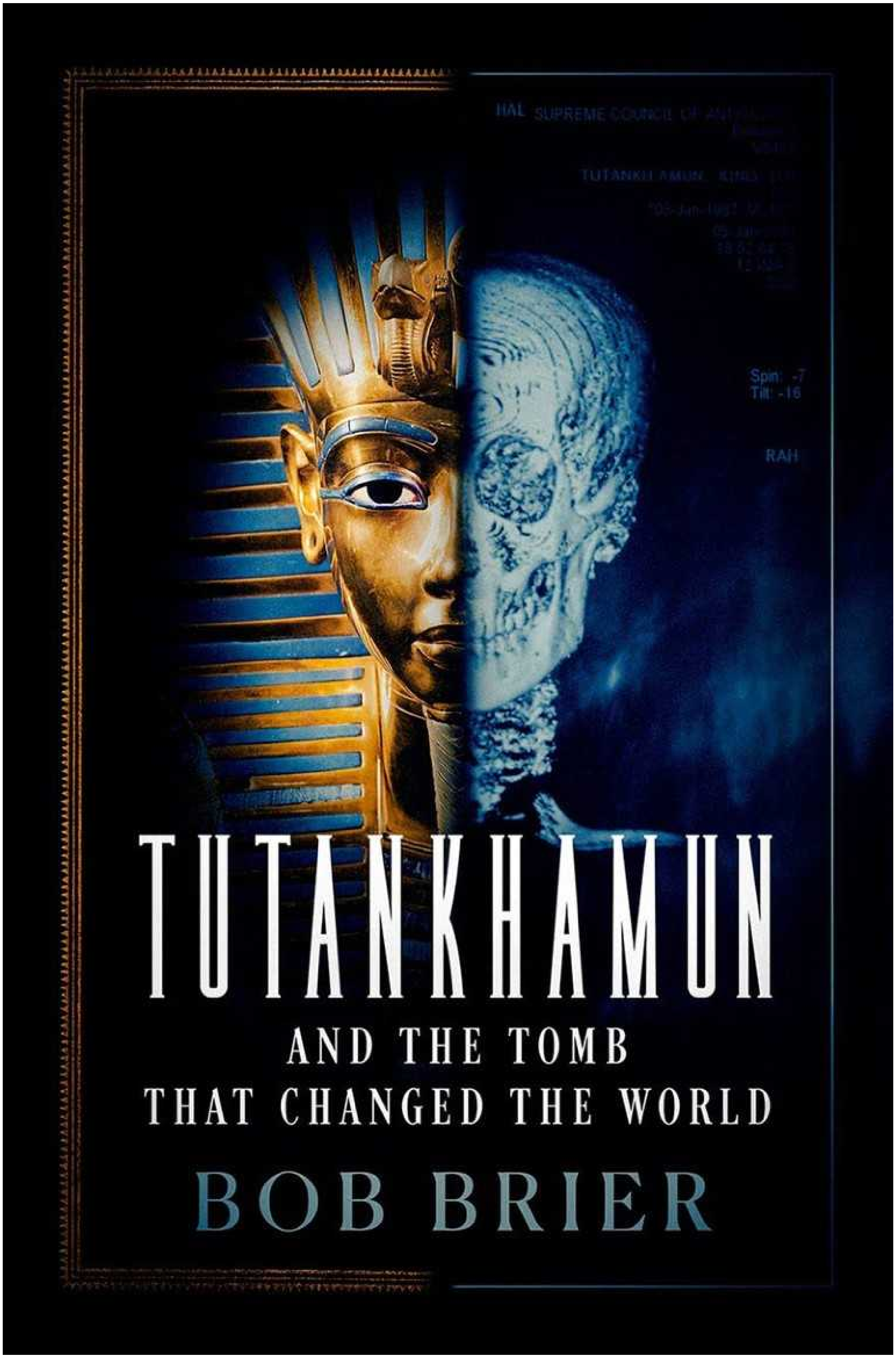
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00643-8

- BOOK REVIEW
- 27 January 2023

The history of profit, and are animals creative? Books in brief

Andrew Robinson reviews five of the best science picks.

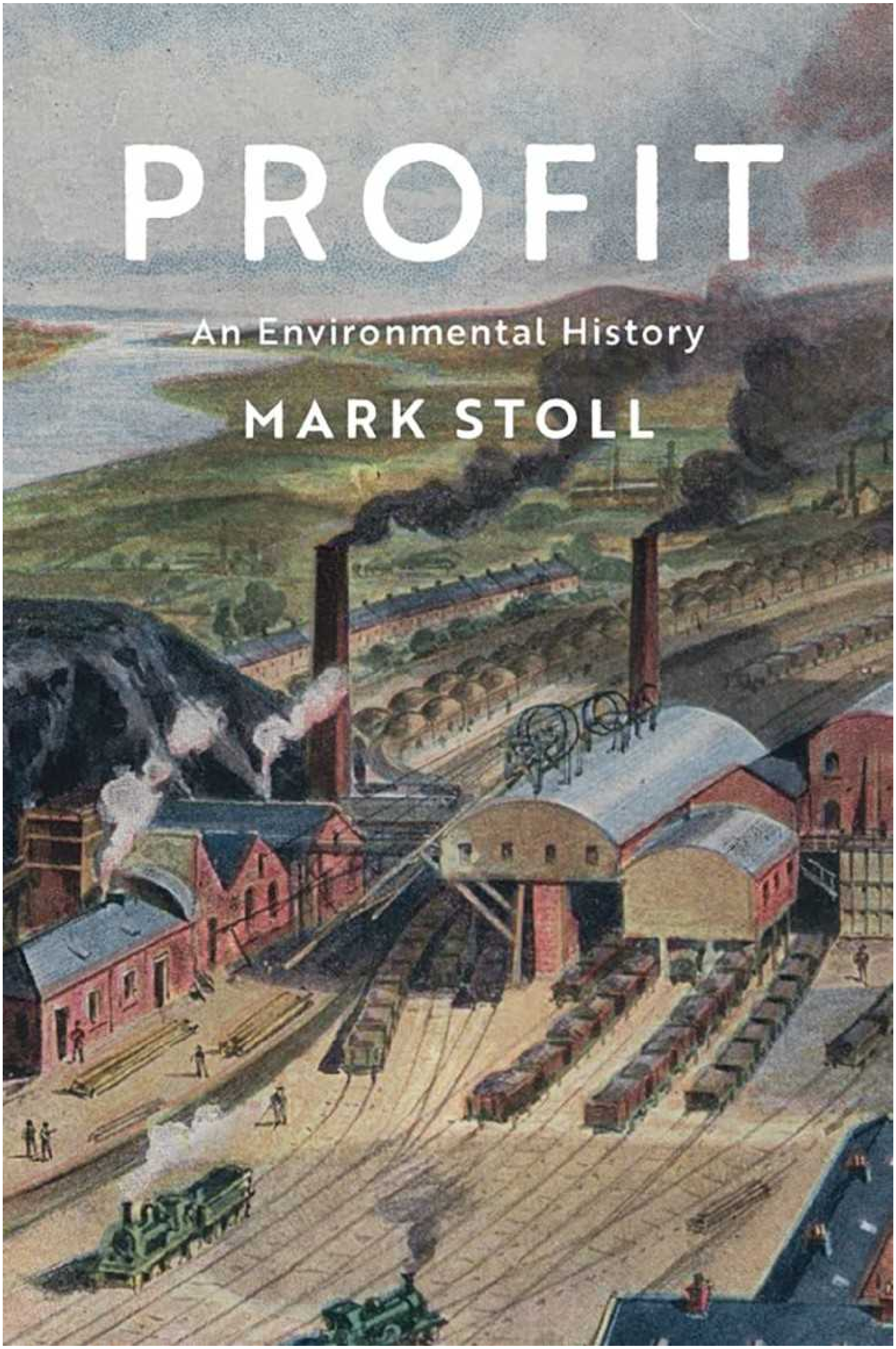
- Andrew Robinson


TUTANKHAMUN
AND THE TOMB
THAT CHANGED THE WORLD
BOB BRIER
Spin: -7
Tilt: -16
RAH

Tutankhamun and the Tomb That Changed the World

Bob Brier *Oxford Univ. Press* (2022)

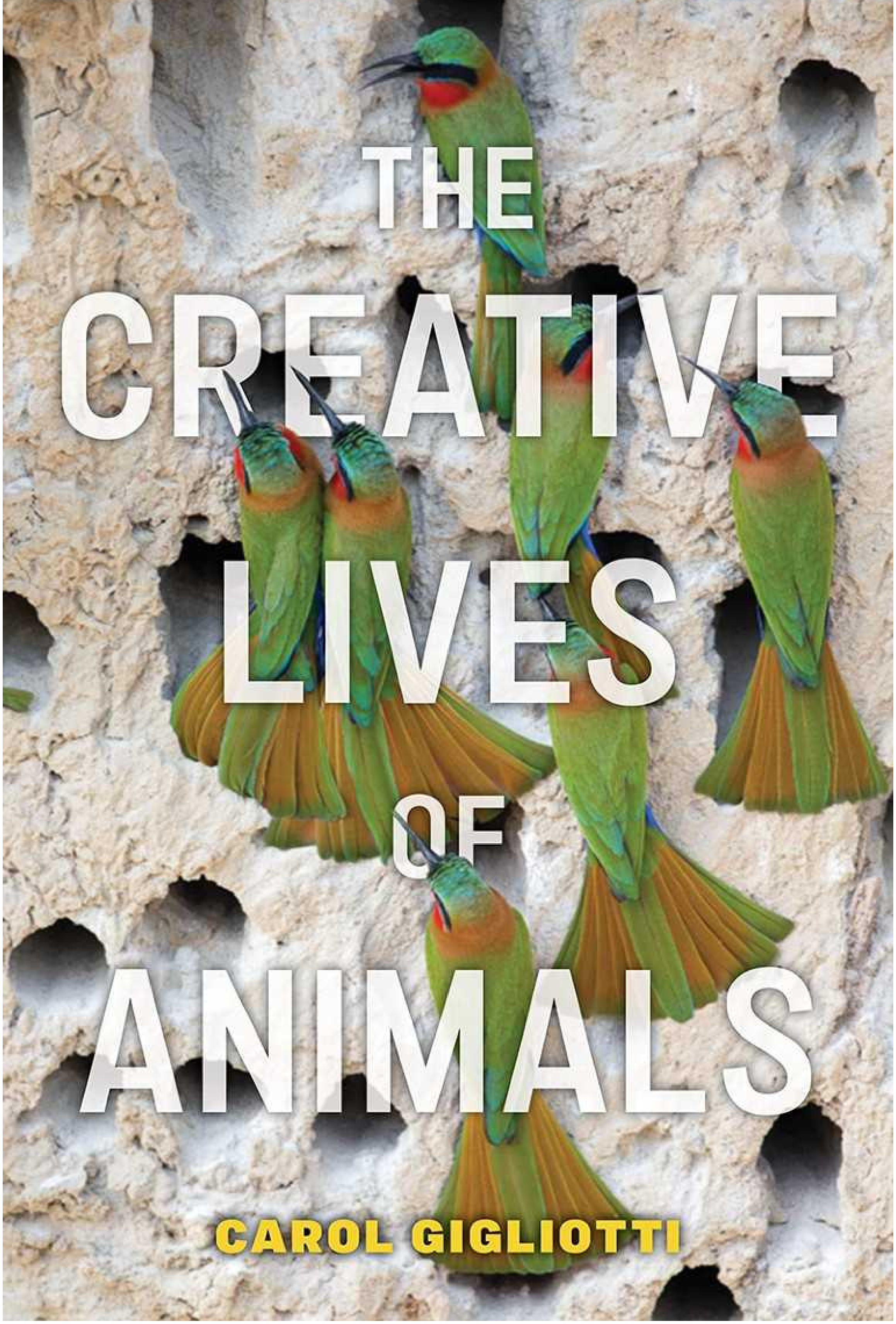
Tutankhamun still raises many questions, argues Egyptologist Bob Brier in his stylish book celebrating the century since the pharaoh's tomb was rediscovered. He has been represented as a frail boy with a club foot, but he might have been an athlete who hunted, and later a warrior, says Brier. As DNA sequencing of mummies improves, Brier anticipates, Tutankhamun's parents will finally be identified, along with diseases that might have caused his death. The second 100 years of Tutankhamun research could be "even more exciting than the first".


PROFIT
An Environmental History
MARK STOLL

Profit

Mark Stoll *Polity* (2023)

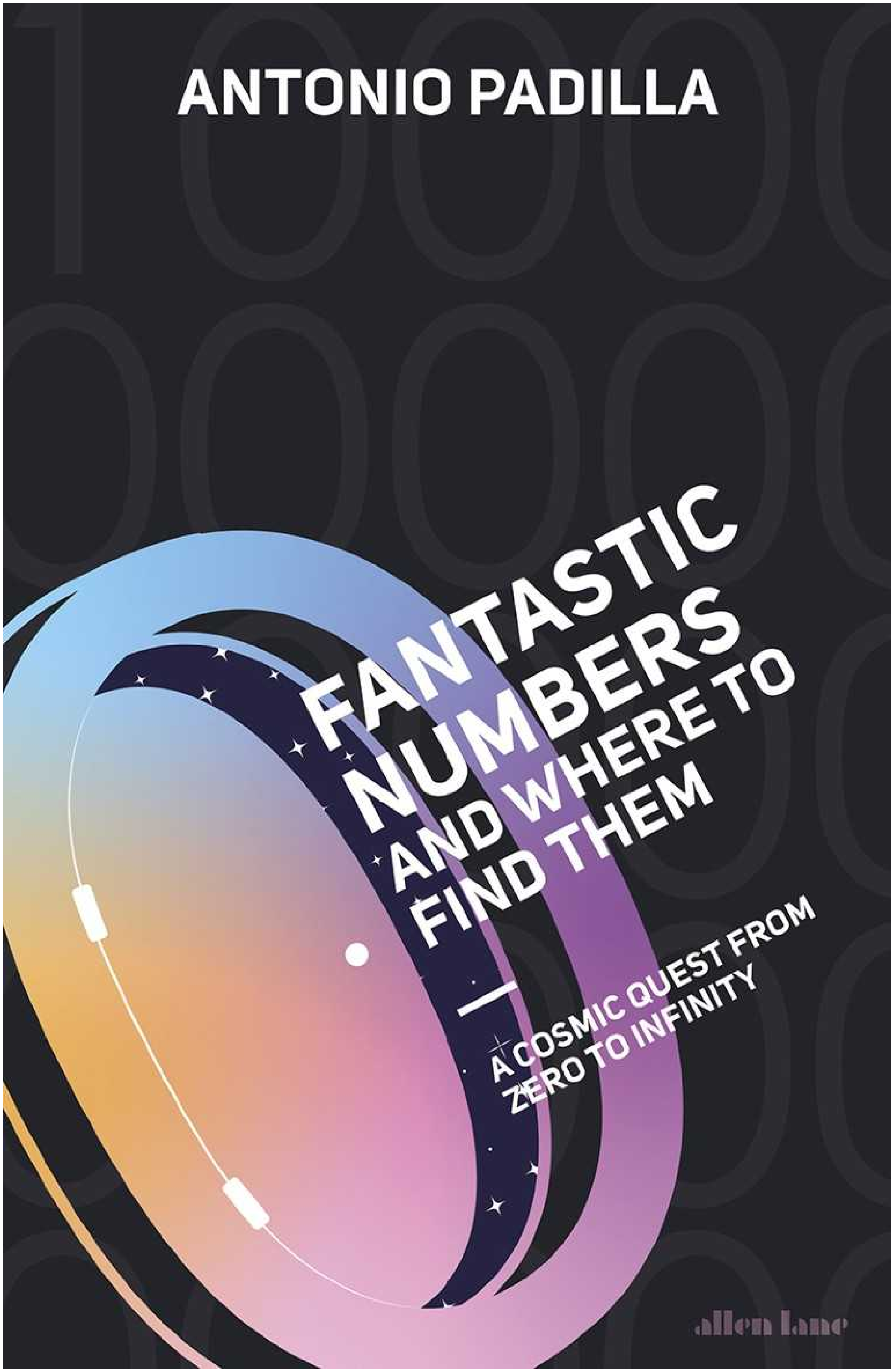
A smartphone is a “Pandora’s box of environmental evils”, writes Mark Stoll. The device epitomizes both the benefits and burdens of capitalism, he argues in his history of profit from ancient times to the present, the first by an environmental historian. With knowledge, skill and stories of inventors, entrepreneurs and conservationists, he traces developments in technology, transportation, energy, communication, trade and finance. We cannot live with capitalism or without it, he says, so we must labour to “ameliorate its worst effects.”


THE CREATIVE LIVES OF ANIMALS
CAROL GIGLIOTTI

The Creative Lives of Animals

Carol Gigliotti *NYU Press* (2022)

Many songbirds are born without the ability to sing. So should those that learn — and other animals — be called creative? Carol Gigliotti interviews scientists who think they should be, and agrees with them. An animal activist, author and artist who has taught design and dynamic media, she defines creativity as a “dynamic process” in which individuals generate “novel and meaningful behaviour” that might affect others at cultural, species and evolutionary levels. Regrettably, her appealing discussion contains no images of animals.


ANTONIO PADILLA
FANTASTIC NUMBERS AND WHERE TO FIND THEM
A COSMIC QUEST FROM ZERO TO INFINITY
allen lane

Fantastic Numbers and Where To Find Them

Antonio Padilla *Allen Lane* (2022)

As a maths student, Antonio Padilla got a zero for a correct proof because it lacked rigour. The experience turned him into a theoretical physicist. “With mathematics, the physicist can dance, and with physics, the mathematician can sing,” he says in his account of numbers that unlock the Universe, such as the cosmological constant. Although not for the mathematically faint-hearted, it sparkles — even comparing quantum theory’s astonishing support of light as a particle rather than a wave with Greta Thunberg suddenly endorsing Donald Trump!

A HISTORY OF THE WIND

ALAIN CORBIN



A History of the Wind

Alain Corbin (transl. William Peniston) *Polity* (2022)

For millennia, little was known about the science of the wind. Aristotle saw it as an elemental fluid, alongside water, earth and fire. In the late eighteenth century, air's chemical composition was revealed; later came the development of meteorology through mapping of global air currents and the concepts of the troposphere, stratosphere and jet stream. All this forms part of Alain Corbin's brief history, but he focuses on the reaction of artists, writers and travellers since ancient Greece to wind's "indecipherable enigma".

Nature **615**, 210 (2023)

doi: https://doi.org/10.1038/d41586-023-00247-2

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00247-2

Opinion

- **Flash floods: why are more of them devastating the world's driest regions?** [07 Mar 2023]
  • Shifting weather, changing settlement patterns and a lack of preparedness mean that dryland areas are most at risk from flooding. Researchers need to focus on data collection, early-warning systems, flood protection and more.

- COMMENT
- 07 March 2023

Flash floods: why are more of them devastating the world's driest regions?

Shifting weather, changing settlement patterns and a lack of preparedness mean that dryland areas are most at risk from flooding. Researchers need to focus on data collection, early-warning systems, flood protection and more.

- Jie Yin,
- Yao Gao,
- Ruishan Chen,
- Dapeng Yu,
- Robert Wilby,
- Nigel Wright,
- Yong Ge,
- Jeremy Bricker,
- Huili Gong &
- …
- Mingfu Guan

Show authors



A family navigates flood waters to reach a village in Mehar, Pakistan, after heavy rains in August 2022. Credit: Akhtar Soomro/Reuters

Last year, around two-thirds of Pakistan was affected by widespread flash flooding, with more than 1,500 people killed and around 33 million made homeless. Almost 2,000 people died in flash floods across Africa, and parts of the United Arab Emirates, Iran, Saudi Arabia, Qatar, Oman and Yemen were inundated with water.

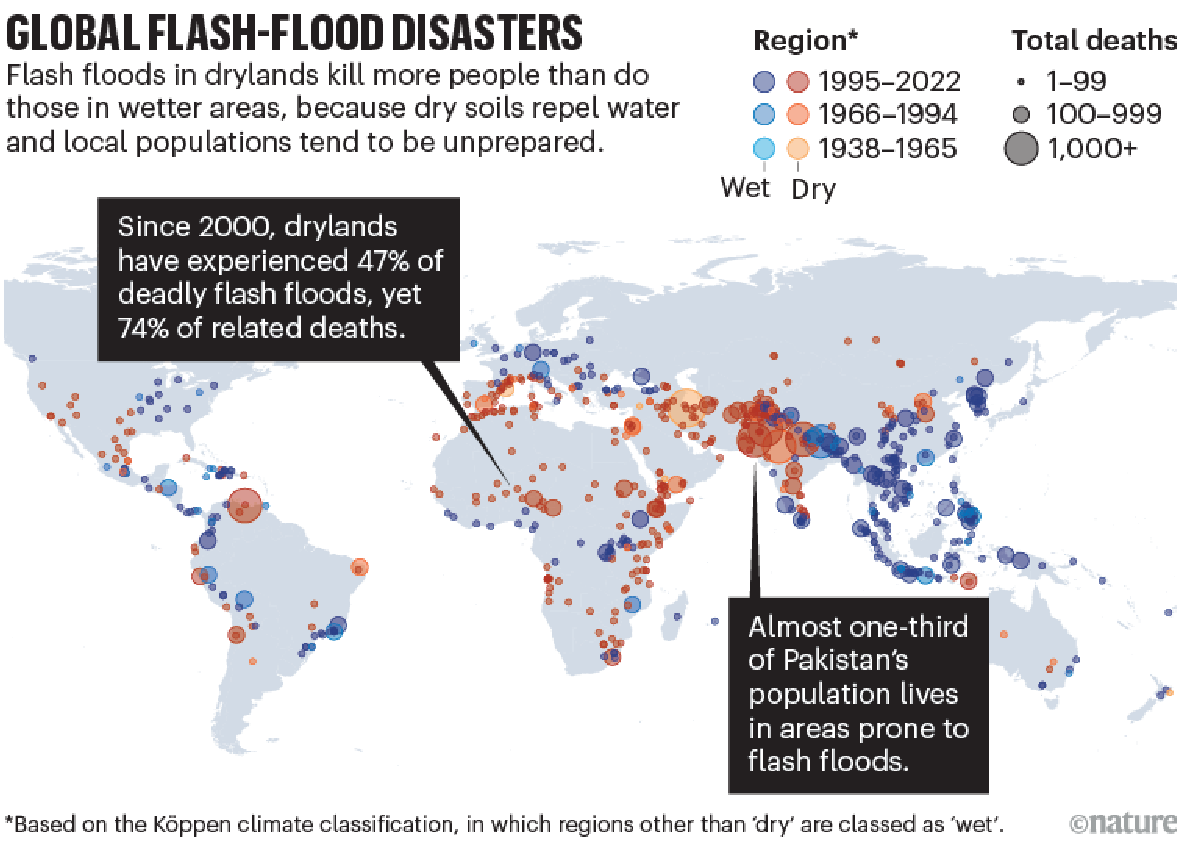
Flash floods are a growing threat in some of the world's driest regions. Deluges can trigger sudden and rapid torrents of run-off that flow down dry river beds and rocky channels.

Because parched soils repel water rather than allowing it to soak in, flash floods can be more devastating in drylands than in wetter areas. Surges can result from relatively small amounts of rain, as little as 10 millimetres in one hour. By comparison, floods in wetter regions typically follow more prolonged bouts of rainfall.

Rapid run-off also erodes soils, adding sediment and debris to the water. The dangers are greatest in mountainous areas, where water can gush suddenly through gullies and down slopes. For example, a flash flood last year in Datong town in Qinghai province, China, washed away more than 1,500 homes in less than one hour.

Yet residents of drylands are underprepared for flash floods, compared with people in the tropics. In dry areas, where drought is a big concern, settlements tend to be near rivers or on floodplains. Dams and levees are low priorities.

Dryland residents are paying a high price. Our analysis using the Emergency Events Database (EM-DAT; www.emdat.be) shows that, since 2000, such regions experienced less than half (47%) of deadly flash floods globally, yet saw almost three-quarters (74%) of related deaths (see ‘Global flash-flood disasters’). The majority of these floods (87%) and associated deaths (97%) occurred in low- and middle-income countries.



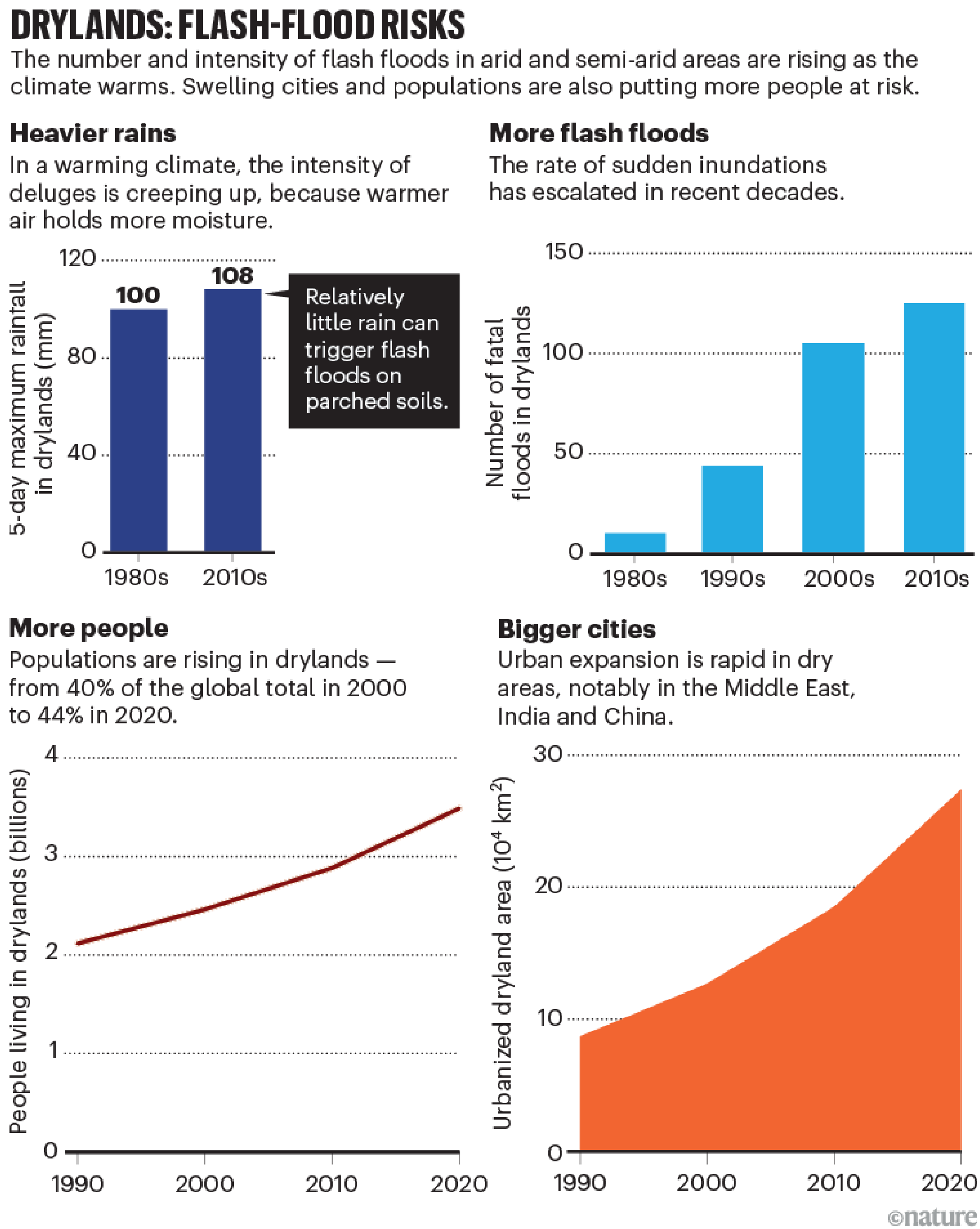
Source: EM-DAT, CRED/UC Louvain (www.emdat.be); Analysis by J. Yin *et al*.

A slew of other factors will put many more people at risk from flash floods in future (see ‘Drylands: flash-flood risks’). Climate change is making such events more intense and frequent[1–3]. In parts of Pakistan, for example, the 5-day maximum level of rainfall is 75% greater today than it was before 1900 (see go.nature.com/41awzzj). Our analysis shows that globally, the rate of dryland flash flooding was 20 times higher between 2000 and 2022 than it was between 1900 and 1999 (numbering 278 and 64 floods, respectively).

More people are living in drylands — from 40% of the global population in 2000 to 44% in 2020, by our calculation. Such regions will host 51% of the world’s population growth by 2025, with half of those people living in low-income countries, and only 1% in developed ones[4]. Cities are expanding rapidly in areas such as the Middle East, India and China[5]. And the proportion of the world’s land area classed as arid or semi-arid is projected to increase as the world warms — from 41% in 2000 to 48% by 2025[4].

Pakistan is one of the most exposed nations. Almost one-third of the country’s population (72 million of 230 million people) lives in high-risk flash-flood zones. As ever, vulnerable groups such as older people, women and children experience the worst effects.

Researchers, practitioners and policymakers need to model, assess and mitigate the risk of flash flooding in drylands in a changing environment. Here, we set out six research priorities for doing so.



Source: EM-DAT, CRED/UC Louvain (www.emdat.be); Analysis by J. Yin *et al*.

Gather data and map the risks

Assessing the likelihood and consequences of flash floods requires a lot of data. These include: meteorological and hydrological measurements; topographic data and digital elevation models; records of assets and previous losses; and census data. However, most dryland regions lack much of this information, because national governments often don't recognize its value and many lack the resources to gather it.

Researchers, international organizations and non-governmental organizations need to help such countries to update their flood-risk information. Initiatives include the Global Flood Partnership run by the European Union (https://gfp.jrc.ec.europa.eu), which is developing flood observation and modelling infrastructure in low- to middle-income countries such as Myanmar. Governments should map flash-flood risks and distribute them online (as the United Kingdom does; see go.nature.com/3xt9hfz) to pinpoint locations that are prone to flooding, restrict construction and target flood-protection measures.

Flood risk needs to be carefully communicated, because it can exacerbate inequalities. In drylands, low-income and minority populations tend to live in flood-prone areas such as riversides[6]. If maps show these areas to be high risk, they could be deprived of local infrastructure investment, increasing their vulnerability. Socio-economic data should therefore also be conveyed to decision makers alongside the maps.

Develop early-warning systems

Early-warning systems for flash floods are operating in some wealthy countries, including the United States and in EU nations. For example, the US system (FLASH) forecasts surface water flows at 10-minute intervals and with a spatial resolution of 1 kilometre. These forecasts combine real-time observations from radar, satellites and sensors with hydrological models, and can provide a few hours' notice of flash floods. If preparation and response plans are also in place, that might give enough time for governments to allocate rescue resources and for residents to protect homes and move to safety points. However, few low- and middle-income nations

have these systems, because they lack the data and models to support them, as well as political incentives and funding.



How climate change and unplanned urban sprawl bring more landslides

To rectify this, the World Meteorological Organization is running a programme to provide weather forecasters and disaster-management agencies with real-time information around flash-flood threats (see go.nature.com/3xsw4fg). So far, 28 warning systems are up and running, with another 39 in development. Eight of the operational systems are in drylands, together with ten of those still in development. More dryland nations need to be engaged.

Forecasting of flash floods in drylands will need tailoring to local contexts, including soil types, terrains and sediment loads. Where data are scarce, uncertainties and limits need to be better understood. Use of wireless sensors and alternative sources of local data, such as use of social media and crowdsourcing, need to be explored.

Boost resilience to sudden floods

Dams, barriers and dikes will need to be installed in drylands to protect settlements. For example, China has been building thousands of kilometres

of levees along the main stream and tributaries of the Yellow River. This is expensive, however, involving hundreds of projects with costs ranging from millions to billions of dollars.

Researchers need to examine the trade-offs. For example, protection measures might encourage more development in flood-prone areas — as was the case in Columbia, South Carolina, which experienced extensive damage from a flash flood in 2015. And dams and levees built to hold back small floods could offer a false sense of security from rare, extreme inundations[7], including one that flooded the Atacama region in South America in 2015. Failure of one protection element, such as a reservoir, can send flood waters cascading downstream, as was narrowly averted at California's Oroville Dam in 2017. The environmental impacts need to be understood, such as those resulting from stopping the silt and sediment that fertilize the floodplain.



Vehicles cross a bridge over the Yolo Bypass in Sacramento, California, where weirs divert flood waters away from houses onto floodplains.Credit:

Fred Greaves/Reuters

Nature-based solutions are cheaper. For example, in Pakistan, the government of Khyber Pakhtunkhwa province has restored almost 350,000 hectares of forests in its ‘Billion Tree Tsunami’ planting drive. In the Yolo Bypass project in Sacramento, California, levees are being replaced with weirs to divert flood waters away from houses and onto natural floodplains. Large residential areas should adopt ‘sponge city’ techniques that reduce run-off, such as permeable pavements, green roofs, rainwater harvesting and swales (ditches). These are being installed in cities throughout China, including Xining in Qinghai province, the largest city on the Tibetan Plateau.

Researchers should aim to quantify how effective natural features are at storing and conveying water, and how resilient they are to extreme floods. Local differences in soil and vegetation conditions must be considered. More needs to be known about how the systems degrade and how to maintain them. Promising projects include the World Bank’s Global Program on Nature-Based Solutions for Climate Resilience and the Sindh Resilience Project in Pakistan, as well as government-led water and soil conservation initiatives in the Loess Plateau in north-central China.

Land-use regulations need to be tightened, such as by preventing informal development in floodplains. Flood-protection standards for buildings need to be improved. For example, the International Residential Code, used as the basis of many national and local guidelines, requires the lowest floors of buildings to be 30 centimetres above the average level of a ‘once a century’ flood. Such levels no longer reflect flood probability in a warming world.

Plan for long-term resettlement

In flood-prone areas, governments should develop plans for temporary and even permanent resettlement. Elevated roads, evacuation routes and emergency shelters must be designed, such as those used in Indonesia for tsunamis and in the United States for hurricanes. In the longer term, as climate-change impacts accelerate, managed retreat might be required[8]. For example, the Chinese government has relocated more than 600,000 people

from flash-flood zones in the arid mountains of southern Shaanxi province over the past decade.

Managed resettlement is more equitable and efficient than leaving people to fend for themselves. Cross-disciplinary research in social science, geography, engineering and psychology is required to investigate the best way to manage retreat, by providing essential services and assistance while avoiding food insecurity and lost livelihoods.

Provide finance beyond disaster aid

Risk-transfer schemes — such as flood insurance, catastrophe bonds and solidarity funds — should be developed for drylands. For example, China is developing a national system of flood insurance products. These have been incorporated into the national emergency-management system and launched in dozens of pilot cities, including Zhengzhou in Henan province. Such products include weather index insurance, which offers compensation for loss of crops and livestock when rainfall exceeds a certain threshold. Catastrophe insurance protects businesses and residences against extreme events that are too expensive for insurers to cover, such as super typhoons or floods that occur only once in 1,000 years.



Five ways to ensure flood-risk research helps the most vulnerable

Algeria has made catastrophe insurance mandatory. Every insurance company there must join its national Catastrophe Risk Insurance Pool to cover disaster losses in proportion to their market share. All properties must be insured, with premiums set by the government.

Researchers need to study the impacts of such schemes, especially on equity. For example, vulnerable families living in flood-prone zones might not be able to afford high premiums[9]. Around 50% of losses in wealthy countries are covered by insurance, but there is less than 10% coverage in many poorer countries. For example, claims for Hurricane Sandy in the United States in 2012 covered 50% of the damages. By contrast, after a record-breaking flash flood in China's Henan province in July 2021, flood insurance claims exceeded US$1.7 billion — only 11% of the total losses.

Partnerships between countries — such as China's 'Belt and Road' project and the 'Build Back Better World' initiative proposed by the G7 group of advanced economies — should be sought to establish solidarity funds that can share the risk of large disasters, as the EU has done. The United Nations climate 'loss and damage' fund, approved last year at the COP27 summit in Sharm El-Sheikh, Egypt, should be leveraged to finance long-term flood-risk reduction in drylands, once mechanisms are formulated.

Improve public awareness and responses

People who live in drylands need to be more aware of flash-flood risks. Policymakers should offer guidance to help residents to adapt their behaviours and take precautionary measures — such as by installing flood-protection structures, knowing evacuation routes and warning signals, identifying areas prone to flooding or landslides, and avoiding walking or driving through flooded areas and standing water.

Researchers need to explore how best to communicate the effectiveness of measures, and how perceptions of risk influence responses to flash floods[10]. Social factors need to be explored — such as age, gender, education, income, ownership, culture, religion and political context. For example, past experience is important for recognizing risks and adapting behaviours. But,

for rare events, memories fade and only catastrophes are remembered for a long time.

Collective behaviours also need to be understood[11]. Social-media messages can provide rapid information about locations, impacts and responses to dryland floods, as revealed through data mining and artificial intelligence. Models can integrate complex human behaviour into quantitative risk assessments, to improve management strategies. Given that flood risks are evolving, it is important to adapt policies as information becomes available. Interdisciplinary analysis will also be needed to evaluate the robustness of flood-resilience strategies for a range of risks and future scenarios[12].

Before more lives are lost, let's offer dryland residents the means to assure their safety.

Nature **615**, 212-215 (2023)

doi: https://doi.org/10.1038/d41586-023-00626-9

References

1. Donat, M. G., Lowry, A. L., Alexander, L. V., O'Gorman, P. A. & Maher, N. *Nature Clim. Change* **6**, 508–513 (2016).

2. Guerreiro, S. B. *et al. Nature Clim. Change* **8**, 803–807 (2018).

3. Fowler, H. J. *et al. Nature Rev. Earth Environ.* **2**, 107–122 (2021).

4. Huang, J., Yu, H., Guan, X., Wang, G. & Guo, R. *Nature Clim. Change* **6**, 166–171 (2016).

5. Ren, Q. *et al. Nature Sustain.* **5**, 869–878 (2022).

6. Rentschler, J., Salhab, M. & Jafino, B. A. *Nature Commun.* **13**, 3527 (2022).

7. Logan, T. M., Guikema, S. D. & Bricker, J. D. *Nature Sustain*. **1**, 526–530 (2018).

8. Hino, M., Field, C. B. & Mach, K. J. *Nature Clim. Change* **7**, 364–370 (2017).

9. Jongman, B. *et al. Nature Clim. Change* **4**, 264–268 (2014).

10. Bubeck, P., Botzen, W. J. W. & Aerts, J. C. J. H. *Risk Anal.* **32**, 1481–1495 (2012).

11. Aerts, J. C. J. H. *et al. Nature Clim. Change* **8**, 193–199 (2018).

12. Dawson, R. J. *et al. Glob. Environ. Change* **21**, 628–646 (2011).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00626-9

Work

- **Chemistry course corrections tackle bias** [06 Mar 2023]
  • The University of York's efforts to decolonize how chemistry is taught holds lessons for other chemistry educators.
- **These tools help visually impaired scientists read data and journals** [06 Mar 2023]
  • Innovative software and modes of presentation are helping to broaden access to the literature.
- **Water-play on the space station isn't just fun and games** [06 Mar 2023]
  • Astronaut Samantha Cristoforetti sloshed fluids around in the International Space Station to gather information for designing fuel tanks.

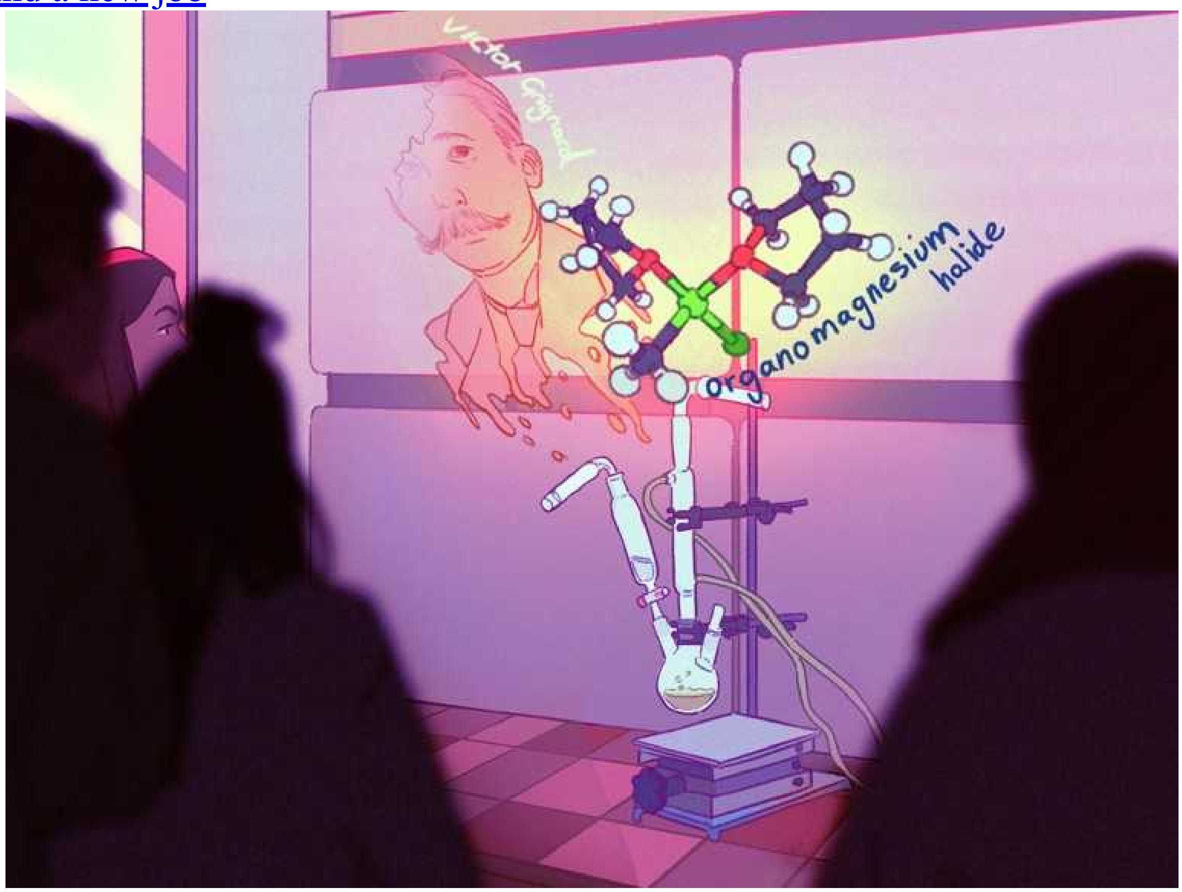
- CAREER FEATURE
- 06 March 2023

Chemistry course corrections tackle bias

The University of York’s efforts to decolonize how chemistry is taught holds lessons for other chemistry educators.

- Katharine Sanderson

Find a new job



Replacing historical chemical names, such as the Grignard reagent, with descriptive ones helps to decolonize chemistry courses. Illustration by Bex Glendining

Flick through a chemistry textbook at school or university, and you'd be forgiven for thinking that chemistry — touted as the central science that connects all other physical-science fields — is the domain of almost exclusively white men. Yet chemistry has origins going back millennia — such as in the papyrus-making procedures of ancient Egyptians, Chinese medicine and distillation methods developed in early Arab cultures. Contributions to modern chemistry are truly global, meaning that low- and middle-income countries can take advantage of chemical advances in areas such as agriculture and medicine as they seek to develop their economies.

To better reflect this breadth, chemistry departments across the globe need to revamp curricula, say those who are pushing for change. It must and will happen, says Avtar Matharu, director of the master's course in green chemistry at the University of York, UK. Matharu, who is Sikh, was born in Nairobi and moved to the United Kingdom as a child. He is also vice-chair of the university's staff race equality forum. "Universities are institutionally racist," he says, but adds that calls for change are too strong for institutions to ignore. "Every university has to respond to the issue," Matharu says.

At the University of York, the chemistry department is an early adopter of a decolonizing agenda and is already known for its progressive stance on inclusion and diversity. This includes LGBT+ friendly policies that take into account the needs of people from sexual and gender minorities; diversity and inclusion training for all undergraduate and postgraduate students; and unconscious-bias training for all recruitment panels. In 2007, it was the first department to achieve a gold Athena Swan award, the highest honour of the UK-wide programme, which recognizes significant gender-equity efforts in the sciences.

In 2019, York students from a variety of disciplines started to urge senior faculty members to decolonize their curricula. This means ensuring that there is a diverse community in departments, that course materials encompass sources from all parts of society and the world, and that courses reflect the lived experience of the diverse student body. In the academic year

2012–13, 21% of York students identified as Black and minority ethnic, rising to 38% in 2022–23.



Decolonizing science toolkit

Within a few months, students on chemistry courses saw that progress was being made in the humanities, but less so in science. So they called on the head of chemistry, Caroline Dessent, to act. She says that the chemistry students' call for action came when the department had already begun to think about how it could address racism and race-related discrimination at all levels. "There was also already quite a desire to do something to improve the culture in the department" for the minority ethnic students and staff, says Dessent, who is white.

In the city of York, 93% of residents are white, according to 2021 census data. "The environment is not very diversified," says Kelechi Uleanya, a postdoctoral research associate in Dessent's research group. Uleanya hails from Nigeria, where she completed her undergraduate and master's studies in chemistry. When she was working as a graduate teaching assistant (GTA) in the undergraduate laboratories at York, she noticed that other GTAs could have benefited from some training in how cultures take different approaches to practical work.

“Some [GTAs] might not have had serious contact with people of different cultures” and therefore lack context about why students might engage in problem sets or labs in different ways, she says. “So, for some of them it’s a shock meeting [students from different cultures] in the labs and their approach to things is totally different because they are coming from another setting.”



Caroline Dessent is head of chemistry at the University of York, UK, and has shaped efforts to decolonize the department’s curriculum in response to student calls for action.Credit: Suzy Harrison

Uleanya joined a steering group, made up of staff and students, that Dessent started in 2020 in response to students’ concerns.

Dessent then wrote to all staff members setting out the decolonizing aims. “It is extremely important when you do equality, diversity and inclusion work that you explain why you’re doing it, because otherwise you’re always at risk of people trying to criticize it,” she says. “Decolonizing is one of

those things that currently does attract a certain amount of negative publicity. Maybe because people just don't understand what they're trying to achieve through it. It's really something very pragmatic."

Distilling lessons

In her newsletter to staff, Dessent laid out why the department was embarking on this course of action. Nationally, there is a 16.1% gap between the number of highest-honours degrees awarded to white UK students compared with Black and minority ethnic UK students, according to 2011 data from the UK National Union of Students (NUS), and this attainment gap is also present at York. Dessent noted that, in the NUS survey, 42% of Black students say that the curriculum does not reflect issues of diversity, equality and discrimination. Ten per cent of transgender students and nearly 25% of women never feel comfortable to speak up in class. "Our undergraduate courses should not disadvantage any student because of their background or characteristics. All of our students should have equal opportunities to thrive in our department," she wrote.

Steps towards achieving this include changing the taught material to be more reflective of contributions from scientists other than those who are white and male, but it also means working to change the culture in the department, she says.

Dessent asked all of her department's faculty members to scrutinize the courses they were currently teaching and find places to introduce some context that would help to diversify the curriculum. For instance, in a course on colloid chemistry that discusses the applications of micelles and gels in cleaning and lubrication, a lecturer incorporated how ancient Babylonians produced one of the first known lubricating gels used as chariot axle grease. In a medicinal chemistry course that covers a global history of medicine, instructors included examples from China, India, the Arab world and Indigenous Australian cultures.



Avtar Matharu helps students identify biases.Credit: Christina Surdhar

Faculty members were also encouraged to include the work of individual non-Western chemists. For example, the medicinal chemistry course highlights the Chinese scientist Tu Youyou, who was the first scientist working in China to receive the Nobel Prize in Physiology or Medicine. She developed the antimalarial drug artemisinin by using her knowledge of traditional Chinese medicine and combining it with modern drug-discovery techniques. "Youyou isolated the active ingredient and did all the medicinal development from that traditional herbal drug to turn it into a pharmaceutical drug," says David Smith, who is white British and a prominent advocate for equity, diversity and inclusion in York's chemistry department. He points out that, ironically, the production of that drug "is now controlled by Western companies, and still isn't available in sub-Saharan Africa", where it is most needed to combat malaria. Teaching this context as part of a pure chemistry course adds a globalized aspect to the

curriculum that shows students how chemistry intersects with traditional knowledge, culture and politics, he says.

These examples, along with many others, have been collated into an evolving document by Uleanya and her colleagues (O. K. Uleanya *et al.* Preprint at ChemRxiv https://doi.org/jzfb; 2022).



How to shrink the gap that holds Black scientists back

Uleanya has been compiling and annotating the resource since October 2021 with funding from the Royal Society of Chemistry in London. “It’s a detailed resource that anybody can go and pull examples from” for teaching high-school or university-level courses, Uleanya explains. When she joined York in August 2016, she was pleased that the department had a reputation for being engaged in equity, diversity and inclusion work, but she still wanted to find more role models. She began to take a deeper look at the literature and found that there were chemists making important discoveries from other countries and other cultures. “People like me contributed a lot,” she says.

And yet in chemistry, lots of the reactions that students learn are named after the often-white male chemist who is credited with discovering or developing them. “There is no need for the name of a reaction to reflect a person,” says Smith. Instead, those reactions can be taught by their reactivity — what

actually happens if that kind of reaction is performed on a certain kind of molecule. For example, the Grignard reaction, which is used in the formation of carbon–carbon bonds, is named after French chemist Victor Grignard. The reaction starts by making a Grignard reagent, a molecule containing a halogen, a magnesium atom and an organic group, which could more descriptively be called an organomagnesium halide, Smith says.

For many students, this is likely to make more sense — removing the need to simply memorize names and instead think more deeply about the chemistry that the reactions exemplify. “Very often, by decolonizing the name you can make things clearer,” says Smith. But wholesale rewriting and reprinting of textbooks, as well as adopting a different name for a historically well-known procedure, is likely to be a slow process, he adds.

Smith emphasizes that the department isn’t trying to dictate to lecturers what they must include in their teaching material. Instead, “it has been a bit more about role modelling the contexts of chemistry”, he says. For instance, lecturers could use examples of chemistry that solves a problem in low-income countries, or of science done by researchers outside the countries that conventionally dominate the literature, he suggests.

Charting prejudices

Matharu isn’t convinced that this kind of collation exercise is ambitious enough. “The easiest thing to do is find lots of examples in the literature for minority ethnic researchers,” he says, but describes this approach as too simple. If a student had a follow-up question, and the teacher or lecturer had included the example as a box-ticking exercise, then the instructor would probably have no further context to add. Matharu has developed his own teaching methods that lead students to realize for themselves how implicitly biased science and the scientific literature are. “I was very keen that we need to start the conversation,” he says.



[Ecologists should create space for a wide range of expertise]

For example, Matharu introduced a section in written essays, in which students have to look up the country and institution of the corresponding authors on all the papers they cite and construct a bar chart. Inevitably, these plots are dominated by US, EU and UK researchers. He first introduced this after marking a student project on coffee production, which predominantly happens in the global south. Matharu totted up the countries of the corresponding authors and realized that there were hardly any from the countries where coffee is actually produced. Now, for all the projects in his postgraduate course on green chemistry, the students must write a narrative about the trend they see in cited authorship.

“That’s really eye-opening,” says Matharu. “When they start to write their narrative, or even see the plot, the plot should already encourage them that their literature is biased.” Matharu now gives points, which count towards a student’s final degree mark, for completing this exercise. “It’s not something that’s too onerous,” he says, “but it’s something that starts to change their behaviour; it gets them to think.”

One student, he says, wrote that the review made them think about the origin of the papers on a deeper level. “That’s quite a profound statement,” he says. “For that particular student, I changed their mindset and changed their way of thinking.” He has dubbed the authorship charts ‘Matharu plots’.

Arinzechukwu Ngoka, who is Black and studied chemistry as an undergraduate in his home country of Nigeria, came to York to pursue the green chemistry master's course. Ngoka says that the graphs are really easy to compile, but that the results are stark. "Matharu plots are an eye-opener to my open bias of Western literatures. I was used to the traditional and biased system thinking that the white scientists dominated the educational system," he says, adding that he now searches for papers from a more diverse range of authors than he did previously.

Matharu hopes that other institutions will take on his idea, and he has presented it to the Royal Society of Chemistry and to some major publishers, he says.



Nazira Karodia calls racism's negative effects on higher education "pervasive" for students.Credit: Edinburgh Napier University

York doesn't stand alone in its decolonization efforts. In her previous roles as pro-vice chancellor with responsibility for regional engagement and as a

science-education professor at the University of Wolverhampton, UK, Nazira Karodia led a range of decolonizing and diversity and inclusion programmes in science and technology. Now deputy vice-chancellor at Edinburgh Napier University, UK, Karodia grew up during the apartheid period of systemic, nationalized racial discrimination in South Africa. "I have had first-hand experience of racism in South African education and society and have seen the pervasive negative influence of racism in higher education in England — a racism that affects so many students' study, lives, expectations and futures," says Karodia.

Karodia says that maintaining a dialogue with stakeholders is key to implementing a successful decolonizing agenda. Students seem to accept change more readily, in her experience, but "staff need to be persuaded, sometimes, by rational argument and an appeal to the values of the liberal institutions. Some staff will never be convinced, and we have to live with that too", she says.

Uleanya has showcased her chemistry-decolonization resource at a number of conferences on equity, diversity and inclusion in science, and interest has been high, she says. A common concern that she hears is that decolonizing the curriculum means a loss of quality. Her advice is to start slowly, including a few new examples. "It has nothing to do with losing the chemistry we have," she says, but rather about including things to make chemistry more global. "Where everybody can feel a sense of belonging. And then to build better chemists."

Nature **615**, 359-361 (2023)

doi: https://doi.org/10.1038/d41586-023-00644-7

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00644-7

- TECHNOLOGY FEATURE
- 06 March 2023

These tools help visually impaired scientists read data and journals

Innovative software and modes of presentation are helping to broaden access to the literature.

- Alla Katsnelson

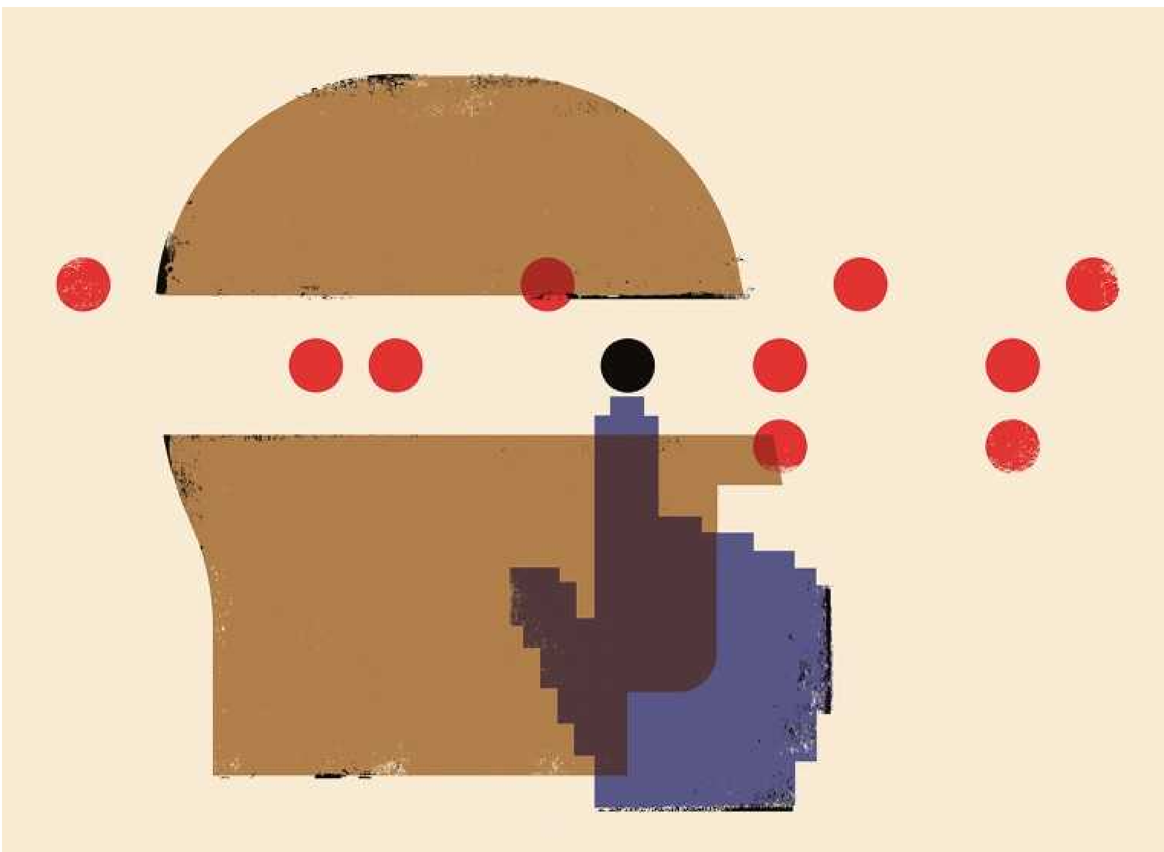


Illustration by The Project Twins

Nearly 240 data scientists have proved themselves nimble enough at their work to be certified instructors for the 'tidyverse', a popular package for manipulating and visualizing data in the R programming language. JooYoung Seo is unique among them — the first blind instructor to gain certification.

Most tidyverse users present data in the form of charts and graphs. Seo, an information and learning scientist at the University of Illinois at Urbana–Champaign, who lost his sight at age ten owing to glaucoma, uses touch, sound and speech. He is one of a small but growing group of researchers working to make science more accessible to people with limited vision. "The overarching challenge is that content is visually designed," Seo says. "But visualization is only one of the representation methods. We can represent data in multimodal ways."

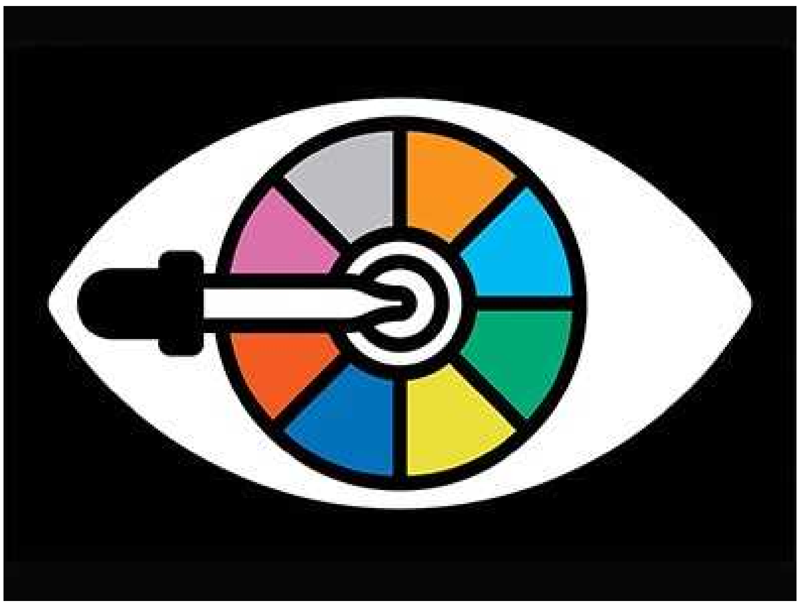
There are no good estimates of how many scientists with low vision are working today, but a 2020 study[1] found that fewer than 100 of 52,124 researchers applying for funding from the US National Institutes of Health (NIH) in 2018 self-identified as having a visual impairment. "It's a fraction of a fraction," says Bonnielin Swenor, an epidemiologist and director of the Disability Research Center at Johns Hopkins University in Baltimore, Maryland, who led that work. That's far short of the total number of scientists with vision disabilities, Swenor adds, because ableist hurdles prevent many scientists from applying for research grants. And it's an even smaller fraction of the number of individuals with visual disabilities overall: in 2017, some 7 million people in the United States (2.17% of the population) were living with 'uncorrectable' loss of visual acuity or with blindness[2]. "If the goal is parity with prevalence in the US — which I argue it should be — we aren't close," she says.

Avoiding workarounds

That's in part because, all too often, even the most basic research activities — accessing the literature, submitting papers, attending conferences and reviewing manuscripts — require time-intensive, personalized workarounds. "Blind scientists probably have a way that works for them," says Daniel Hajas, an innovation manager at the Global Disability Innovation Hub at

University College London who works to accelerate assistive technologies to market. “But it’s not generalizable and it took them a long time to get there.”

Swenor, who has degenerative retinal disease, recalls that a mentor early in her career advised her that she would have to work four times as hard as sighted colleagues to succeed in science. “That should not be the path, for anyone,” she says. But awareness of the need for accessibility in science is growing, and new tools are on the way.



Colour me better: fixing figures for colour blindness

People whose visual acuity is too low for them to read magnified text generally use screen readers, which convert digital text into either synthesized speech or a tactile (Braille) display.

Screen readers navigate digital documents using tags that identify elements such as headers, figures, tables and footnotes. But the devices often stumble on PDF files, which have long been the default digital format for journal articles and other research materials, because they generally lack such tags. The two-column format widely used in journals can also confuse screen readers, which generally read from left to right across a page. PDFs can be made accessible, but it’s challenging to do and many publishers haven’t got there yet. “Improving PDF accessibility requires changes to culture, systems

and processes that can be challenging for publishers to achieve," says Jude Robinson, a global lead for Springer Nature Digital in London. "We are committed, however, to doing this." (*Nature*'s news team is editorially independent of its publisher, Springer Nature.)

An online tool called SciA11y, developed by the Allen Institute for Artificial Intelligence (AI2) in Seattle, Washington, uses machine learning to extract the content and structure of a PDF (and other file formats, such as LaTeX) and re-render it in HTML, creating a table of contents containing links to tagged section headings that can be navigated with screen readers. AI2 has also built in functionality such as bidirectional navigation between in-line citations and their corresponding references in the bibliography, says Jonathan Bragg, who co-leads the project.

In a detailed survey[3], six scientists with vision loss described how they frequently found themselves unable to access or read PDFs. One respondent mentioned that they encountered problems two-thirds of the time, and that they use at least six different approaches to read papers. "It was eye-opening to see the range of tools that people use when reading — and the struggles they have when those papers have not been formulated appropriately," Bragg says.



Speaking in code: how to program by voice

At the moment, SciA11y — ‘a11y’ is Internet shorthand for accessibility — is an online demo: researchers can upload PDFs to re-render them in HTML. But the team is still working on key functionality, Bragg says. For example, the software still makes mistakes, sometimes failing to pull out headings and leaving them as body text in the HTML. It also struggles with tables and images, which is a current focus of development.

In that respect, SciA11y is not unique: for visually impaired people, tables and images present challenges that are even harder to solve than text itself. Precious little scientific literature — whether in PDF, HTML or another format — includes textual descriptions of figures, called alt text, that would enable a blind or low-vision person to understand the images. What’s more, it’s not always clear what sort of explanation would be most relevant. “There are different schools of thought,” says Amy Bower, a blind physical oceanographer at the Woods Hole Oceanographic Institution in Massachusetts. “Should you just describe what’s there, or should you add the interpretation?”

Some researchers, including Bower, solve this issue by working with sighted interpreters. Hajas is collaborating with researchers at the Massachusetts Institute of Technology in Cambridge and others to design Olli, a screen-reader interface that allows users to navigate up and down different levels of description — from a single sentence explaining a figure’s main takeaway, to a more detailed characterization of the axes and legend, to the actual data values from which it is built. Olli currently supports basic chart types, such as bar charts and scatterplots, and Hajas says that the team is working on new features, such as heat maps.

Seo is also developing a data visualization tool, and he is drawing on more than just speech. The tool, called the multimodal access and interactive data representation system (MAIDR), encodes data as both sounds — termed sonification — and Braille, providing tactile analysis with the help of a refreshable Braille display. “I can hear the trend of the data in sound, and feel the pattern,” Seo says.

Experiencing data through touch can be especially powerful, says Mona Minkara, a computational chemist and bioengineer at Northeastern University in Boston, Massachusetts, who began to lose her vision as a

young child. Minkara collaborated on a 2022 study[4] that described representing data as 3D-printed graphics called lithophanes. Produced from plastic thin enough for light to shine through, lithophanes can encode multiple forms of chemical and life-science data — for example, a scanning electron micrograph of a butterfly wing, the bands of an electrophoresis gel or the ultraviolet spectrum of a protein. The technology, she says, allows her and her sighted laboratory colleagues to engage with the same data at the same time.



NatureTech hub

Feeling a change in protein function through your fingers as a difference in thickness, for example, "adds a layer of knowledge", she says, regardless of whether a researcher can perceive that change visually. "It's going to embed into your imagination and be more a part of you, so you understand it on a deeper level."

Sonification provides another way for scientists with visual limitations to 'see' data, for instance in astronomy. "We are used to thinking that astronomy is a visual science, but most of the data that we get are just numbers," says Anita Zanella, an astronomer at the Italian National Institute for Astrophysics in Rome. "We translate them into images because that's a way for us to make sense of them", but ears are often better at detecting faint signals. By translating numerical values into sounds with certain parameters

— for example, a star's brightness might be encoded as pitch — researchers can home in on important changes[5].

Researchers in genomics and geology are also exploring sonification, although shared principles and standards that would allow scientists to compare data in auditory formats are still in development, Zanella says. Still, simple online tools, such as the Highcharts Sonification Studio, developed at the Georgia Institute of Technology in Atlanta, allow researchers in any field to upload data and explore ways to represent them aurally.

In addition to new tools, mindsets have advanced substantially over the past five years as researchers and developers increasingly consider accessibility, according to several researchers interviewed for this article. And in the United States, at least some federal funding bodies have made accessibility a priority. On 30 December, the NIH published recommendations (see go.nature.com/3gtmtk2) for breaking down ableism and creating equitable and inclusive research practices. (Swenor co-chaired the NIH's Subgroup on Individuals with Disabilities, which developed the proposals.)

But for researchers with vision loss to truly succeed, tools will have to evolve beyond piecemeal efforts, Hajas says. "Unless we make the whole ecosystem accessible — how to find articles, how to read the text, how to access the diagrams and everything else — there will be missing pieces that essentially disable people from becoming good scientists," he says.

Nature **615**, 362-363 (2023)

doi: https://doi.org/10.1038/d41586-023-00645-6

References

1. Swenor, B. K., Munoz, B. & Meeks, L. M. *PLoS ONE* **15**, e0228686 (2020).
2. Flaxman, A. D. *et al. JAMA Ophthalmol.* **139**, 717–723 (2021).
3. Wang, L. L. *et al.* Preprint at https://arxiv.org/abs/2105.00076 (2021).

4. Koone, J. C. *et al. Sci. Adv.* **8**, eabq2640 (2022).

5. Zanella, A. *et al. Nature Astron.* **6**, 1241–1248 (2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00645-6

| Section menu | Main menu |

- WHERE I WORK
- 06 March 2023

Water-play on the space station isn’t just fun and games

Astronaut Samantha Cristoforetti sloshed fluids around in the International Space Station to gather information for designing fuel tanks.

- Linda Nordling



Samantha Cristoforetti is a European Space Agency astronaut who lives in Cologne, Germany. Credit: ESA/NASA

I’m an astronaut with the European Space Agency. Last year, I spent five months on the International Space Station (ISS) — from late April to mid-October — and spent the last month as station commander. Before returning to Earth, my team and I set aside time for a water ‘play date’. Here, inside the ISS, I am demonstrating how water behaves in zero gravity.

You can use little tricks to make sure the water stays where you want it. Surface tension keeps the water bubble together, and you can move it by using a straw to pull on it gently, or by blowing on it. If the bubble is small enough, you can drink it. We recycle all the water onboard.

Weightlessness is not just exhilarating, it’s also an opportunity to study fundamental physics. There’s a lot of research on the space station on fluid dynamics. One study I was personally involved in looked at the sloshing behaviour of different kinds of fluid and mixes of fluids and gases in containers. The results are important for designing fuel tanks, especially for use in space, but also on Earth.

The photo was taken in the Japanese Experiment Module. This is the largest single ISS module and the least cramped, so we often use it for talking to the media or to schoolchildren. When we communicate with them, we use props, such as the balls behind me, which are models of planets and the Moon. The round thing behind me is the module’s airlock. We use it for deploying satellites, as well as hardware such as scanners for science experiments.

This was my second time on the ISS. I adapt quickly to space — I really enjoy that feeling of weightlessness. It’s a lot harder for me to come back to Earth.

I don’t know when I’m going up there again — or if I will. We’ll see how the US-led Artemis programme — which will return people to the Moon in the coming decade — evolves. Maybe there will be an opportunity for me.

Nature **615**, 366 (2023)

doi: https://doi.org/10.1038/d41586-023-00646-5

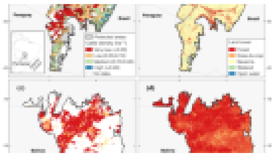
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00646-5

Research

- **Tropical deforestation causes large reductions in observed precipitation** [01 Mar 2023]
  • A pan-tropical analysis using satellite, station-based and reanalysis datasets shows that deforestation causes reduced precipitation, and demonstrates that the effect increases with spatial scale.
- **Coastal phytoplankton blooms expand and intensify in the 21st century** [01 Mar 2023]
  • Satellite observations reveal global increases in the extent and frequency of phytoplankton blooms between 2003 and 2020 and provide insights into the relationship between blooms, ocean circulation and sea surface temperature.
- **Evolution of the germline mutation rate across vertebrates** [01 Mar 2023]
  • Using sequencing and comparing high-coverage genomes, the germline mutation rates across vertebrates are quantified.
- **Cardiogenic control of affective behavioural state** [01 Mar 2023]
  • Direct elevation of heart rate using noninvasive optogenetics in mice influences anxiety-like behaviours in specific environmental contexts, and the posterior insular cortex is implicated in this integration of signals from the heart with environmental risk information.
- **Coordination of bacterial cell wall and outer membrane biosynthesis** [01 Mar 2023]
  • A study demonstrates that specific interactions between the two committed enzymes for the synthesis of lipopolysaccharide and peptidoglycan enable coordinated assembly of the outer membrane and cell wall in the Gram-negative pathogen Pseudomonas aeruginosa.
- **Autoimmunity in Down’s syndrome via cytokines, CD4 T cells and CD11c+ B cells** [22 Feb 2023]
  • An autoimmune-prone state of steady-state cytokinopathy, hyperactivated CD4 T cells and ongoing B cell activation contributes to a breach in immune tolerance in individuals with Down’s syndrome.
- **H3K4me3 regulates RNA polymerase II promoter-proximal pause-release** [01 Mar 2023]
  • Acute loss of H3K4me3 does not have detectable effects on transcriptional initiation, but leads to a widespread decrease in transcriptional output, an increase in RNA polymerase II

pausing and slower elongation

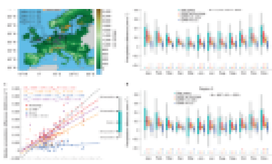
- **Multivariate BWAS can be replicable with moderate sample sizes** [08 Mar 2023]
  -
- **Reply to: Multivariate BWAS can be replicable with moderate sample sizes** [08 Mar 2023]
  -



Changes in land use enhance the sensitivity of tropical ecosystems to fire-climate extremes

19 January 2022

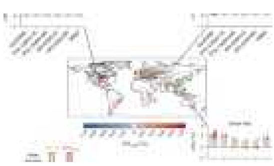
Sujay Kumar, Augusto Getirana, … Niels Andela



Empirical estimate of forestation-induced precipitation changes in Europe

05 July 2021

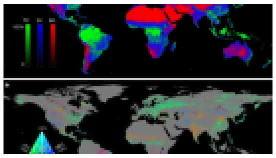
Ronny Meier, Jonas Schwaab, … Edouard L. Davin



Historical deforestation locally increased the intensity of hot days in northern mid-latitudes

23 April 2018

Quentin Lejeune, Edouard L. Davin, … Sonia I. Seneviratne



[Global land change from 1982 to 2016]

08 August 2018

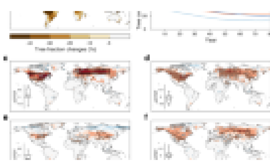
Xiao-Peng Song, Matthew C. Hansen, … John R. Townshend



[Hysteresis of tropical forests in the 21st century]

05 October 2020

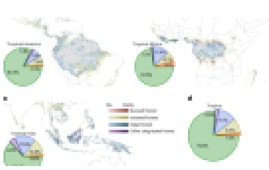
Arie Staal, Ingo Fetzer, … Obbe A. Tuinenburg



[Deforestation intensifies daily temperature variability in the northern extratropics]

10 October 2022

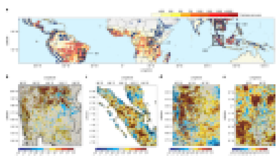
Jun Ge, Qi Liu, … Weidong Guo



[**Comparable biophysical and biogeochemical feedbacks on warming from tropical moist forest degradation**]

02 March 2023

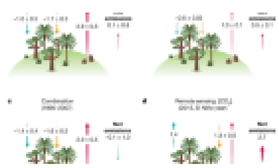
Lei Zhu, Wei Li, … Jingmeng Wang



[**Deforestation-induced warming over tropical mountain regions regulated by elevation**]

14 December 2020

Zhenzhong Zeng, Dashan Wang, … Eric F. Wood



[**The tropical forest carbon cycle and climate change**]

25 July 2018

Edward T. A. Mitchard

- Article
- Open Access
- Published: 01 March 2023

Tropical deforestation causes large reductions in observed precipitation

- C. Smith ,
- J. C. A. Baker &
- D. V. Spracklen

Nature **volume 615**, pages 270–275 (2023)

- 12k Accesses
- 673 Altmetric
- Metrics details

Abstract

Tropical forests play a critical role in the hydrological cycle and can influence local and regional precipitation[1]. Previous work has assessed the impacts of tropical deforestation on precipitation, but these efforts have been largely limited to case studies[2]. A wider analysis of interactions between deforestation and precipitation—and especially how any such interactions might vary across spatial scales—is lacking. Here we show reduced precipitation over deforested regions across the tropics. Our results arise from a pan-tropical assessment of the impacts of 2003–2017 forest loss on precipitation using satellite, station-based and reanalysis datasets. The effect of deforestation on precipitation increased at larger scales, with satellite datasets showing that forest loss caused robust reductions in precipitation at scales greater than 50 km. The greatest declines in precipitation occurred at 200 km, the largest scale we explored, for which 1 percentage point of forest loss reduced precipitation by 0.25 ± 0.1 mm per month. Reanalysis and station-based products disagree on the direction of precipitation responses to forest loss, which we attribute to sparse in situ tropical measurements. We estimate that future deforestation in the Congo will reduce local precipitation by 8–10% in 2100. Our findings provide a compelling argument for tropical forest conservation to support regional climate resilience.

Main

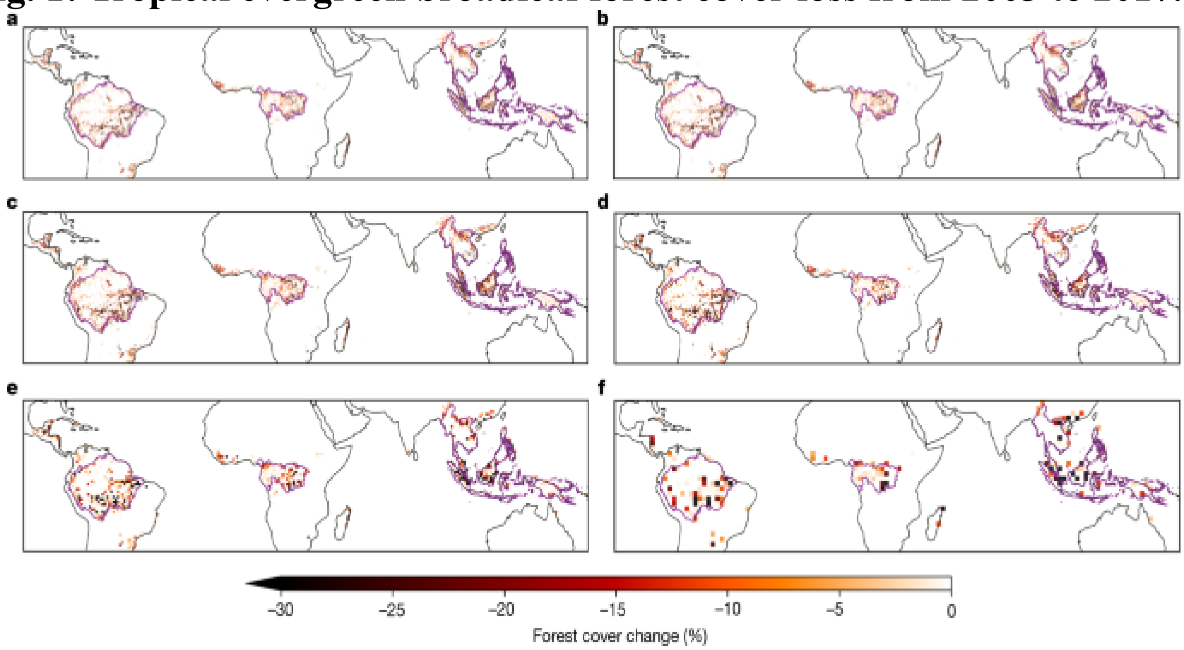
Tropical forests play an important role in moderating local, regional and global climate through their impact on energy, water and carbon cycles[3]. Crucially, tropical forests control local-to-regional rainfall patterns[1,2]. Evapotranspiration from tropical forests is a strong driver of regional precipitation[4,5] contributing up to 41% of basin mean rainfall over the Amazon and up to 50% over the Congo[6]. Evergreen tropical forests are dependent on high annual rainfall for their survival and productivity[7], and forest–rainfall feedbacks have been highlighted as an important determinant of tropical forest stability[4,5,8], amid concerns that the exacerbating impacts of droughts and deforestation could threaten their viability[9].

Rapid loss of forests is occurring across the tropics[10]. Tropical deforestation warms the climate at local-to-global scales by changing the surface energy balance and through emissions of carbon dioxide[3]. The impact of tropical deforestation on precipitation is less certain with a range of processes operating at different scales. Small-scale deforestation over the southern Amazon has been shown to increase precipitation frequency[11,12] owing to thermally[13] and dynamically[12] induced circulations. At larger scales, deforestation reduces precipitation recycling leading to a reduction in precipitation[1,14]. Over Indonesia, deforestation has been linked to declining precipitation[15], and exacerbation of El Niño impacts[16]. Global and regional climate models predict annual precipitation declines of 8.1 ± 1.4% for large-scale Amazonian deforestation by 2050 (ref. [17]), but an observational study of the impacts of tropical deforestation on precipitation across spatial scales is lacking.

Here we present a pan-tropical assessment of the impact of forest loss on precipitation based on measurements. We use a satellite dataset of forest cover change over the period 2003–2017 to identify areas of forest loss, with a focus on evergreen broadleaf forests of the Amazon, Congo and Southeast Asia (SEA; Fig. 1). To provide a robust assessment of the impacts of deforestation on precipitation, we analysed 18 different precipitation datasets, including satellite ($n = 10$), station-based ($n = 4$) and reanalysis ($n = 4$) products (Extended Data Table 1). We compared the precipitation change over pixels experiencing forest loss with neighbouring pixels that had experienced less forest loss (Methods). This comparison against

neighbouring pixels that will have experienced similar climate change focuses our analysis on the changes due to forest loss. To explore the impact of forest loss across scales, we analysed the impacts of forest loss on coincident precipitation at a series of spatial resolutions ranging from roughly 5 km to 200 km (0.05°, 0.1°, 0.25°, 0.5°, 1.0° and 2.0°).

Fig. 1: Tropical evergreen broadleaf forest cover loss from 2003 to 2017.



a–**f**, Forest cover change at 0.05° (**a**), 0.1° (**b**), 0.25° (**c**), 0.5° (**d**), 1.0° (**e**) and 2.0° (**f**) resolution. The Amazon Basin, Congo Basin and SEA regions used in this study are outlined in purple. Maps of the different regions generated using Cartopy and Natural Earth[51]. Forest loss data from ref. [10].

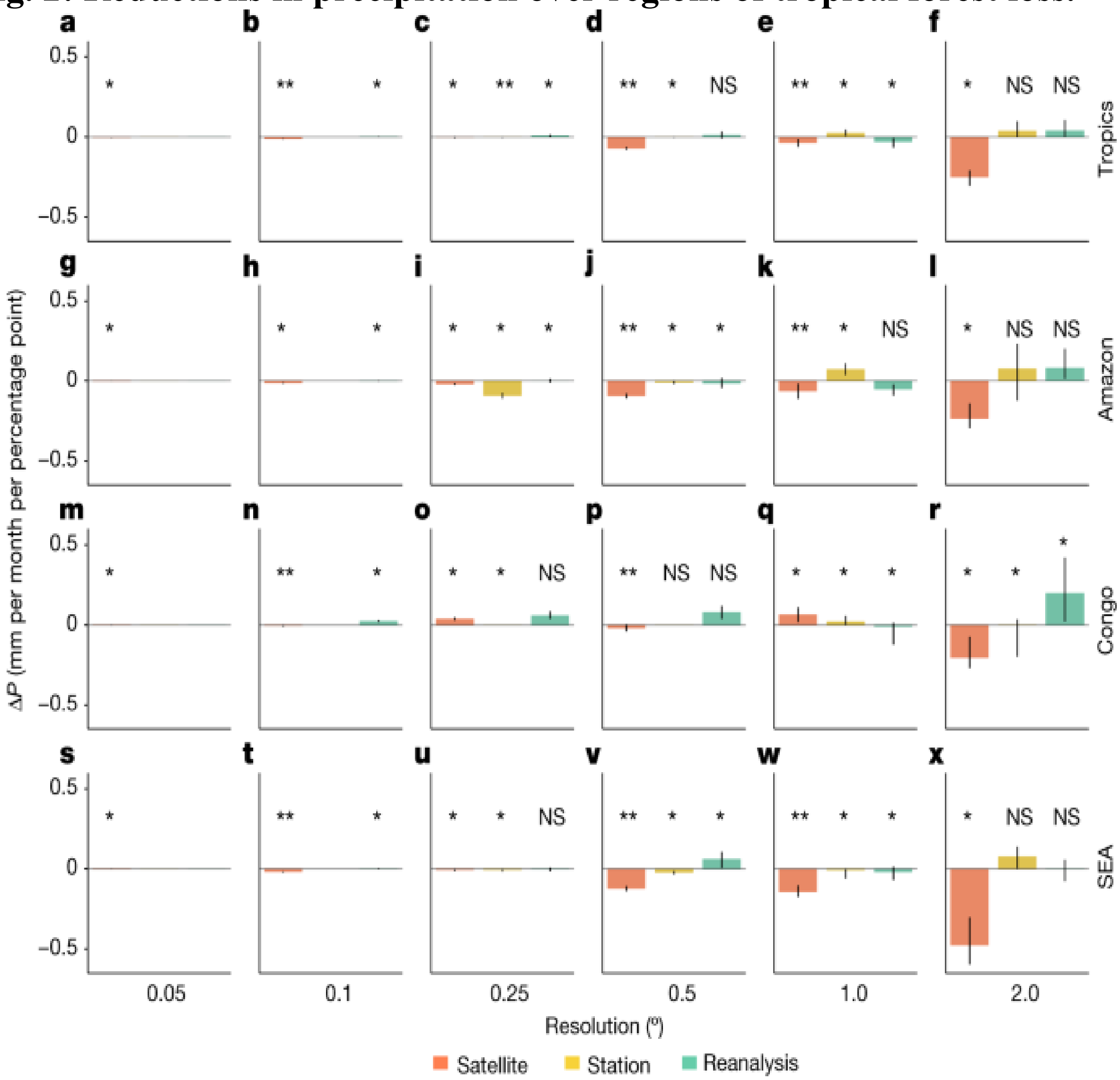
Source Data

Precipitation response to forest loss

Observed precipitation responses to tropical forest loss across multiple spatial scales and precipitation products are presented in Fig. 2. Satellite-based precipitation datasets suggest that tropical forest loss causes statistically significant ($P < 0.05$) declines in median annual mean precipitation at all scales analysed. At larger scales (>0.5°), reductions

exceed 0.03 mm per month for each percentage point loss of forest cover (Fig. 2d–f). The largest changes are observed at the 2.0° scale (approximately 220 km at the Equator; Fig. 2f), for which each percentage point reduction in forest cover causes 0.25 ± 0.1 mm per month reduction in annual precipitation.

Fig. 2: Reductions in precipitation over regions of tropical forest loss.



a–**r**, Bars indicate the median absolute change in annual precipitation (millimetres per month) per percentage point of forest loss over 2003 to 2017 in each region (tropics (**a**–**f**), Amazon (**g**–**l**), Congo (**m**–**r**), SEA (**s**–**x**)) for each precipitation dataset category (satellite, station and reanalysis). Results are shown for forest loss scales of 0.05° (**a**,**g**,**m**,**s**), 0.1° (**b**,**h**,**n**,**t**),

0.25° (**c**,**i**,**o**,**u**), 0.5° (**d**,**j**,**p**,**v**), 1.0° (**e**,**k**,**q**,**w**) and 2.0° (**f**,**l**,**r**,**x**). Statistically significant (*$P < 0.05$; **$P < 0.01$) and nonsignificant (NS) differences in changes in mean precipitation (calculated as a multi-annual mean over 2003–2007 compared with 2013–2017) over deforested regions compared with control regions are indicated. Error bars show ±1 standard error from the mean. Datasets used in this analysis are detailed in Extended Data Table 1. ΔP, precipitation change.

Source Data

Analysis of precipitation change as a function of forest loss confirms larger reductions in precipitation for larger reductions in forest cover (Extended Data Fig. 1), although with considerable variability, as seen in the modelled response[18]. Observed reductions in precipitation are consistent across satellite datasets, with all ten satellite precipitation products agreeing on the sign of the rainfall response at 2° over the tropics (Extended Data Fig. 2). At the 2° scale, significant ($P < 0.05$) reductions in annual mean precipitation with forest loss were observed across all tropical regions (Fig. 2). Reductions in precipitation at 2° based on satellite datasets ranged from 0.48 ± 0.36 mm per month in SEA to 0.23 ± 0.12 mm per month in the Amazon, and 0.21 ± 0.19 mm per month in the Congo for each percentage point loss in forest cover, with at least 8 out of 10 satellite datasets agreeing on the sign of the response within each region (Extended Data Fig. 2). In SEA, it has been suggested that proximity to the ocean and the replacement of tropical forest with plantations as opposed to pasture or cropland may cause reduced sensitivity of precipitation to deforestation[1]. Our analysis suggests that forest loss in SEA causes reductions in precipitation consistent with or greater than reductions in precipitation in the Amazon and Congo.

Station-based datasets and reanalysis products exhibit contrasting annual mean precipitation responses to deforestation at 2.0° (Fig. 2). Across the tropics, station-based and reanalysis datasets showed no statistically significant changes in annual mean precipitation due to forest loss (Fig. 2f), and there was little agreement with satellite datasets at the regional scale (Fig. 2l,r,x), with some non-satellite precipitation products showing small increases in annual mean precipitation due to forest loss. Sparse in situ measurements across the tropics, particularly over regions of forest loss, mean that station-based datasets provide a weak constraint on precipitation

changes. A comparison of station-based precipitation datasets revealed higher levels of uncertainty in the tropics, including the Amazon[19]. In regions of sparse data such as tropical forests[20], interpolation methods may mask precipitation changes driven by forest loss. Reanalysis products, which are numerical models constrained by empirical data, are also expected to be less reliable in regions where in situ data are limited[21]. Our results indicate that precipitation data based on satellite remote-sensing measurements may have an advantage over tropical forest regions where in situ measurements are sparse or unavailable. For these reasons, we focus our analysis on satellite-based datasets and identify where agreement between datasets exists.

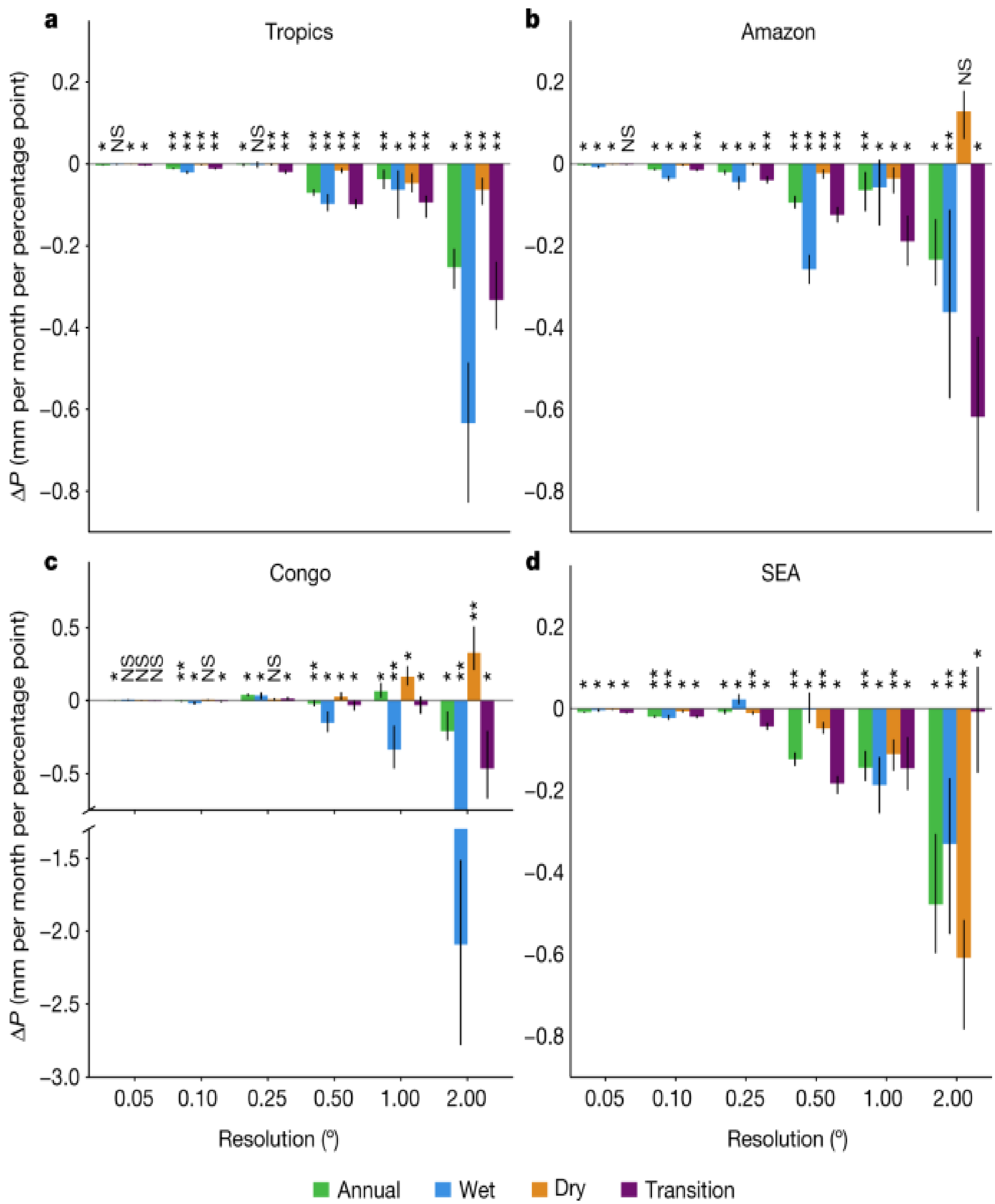
Our results are robust (Extended Data Fig. 3) to a range of methodological assumptions including the length of analysis period, the choice of start and end period and the spatial extent of control pixels (Methods). Our analysis period includes the 2015–2016 El Niño that resulted in negative precipitation anomalies over many tropical land regions (Supplementary Fig. 1). We found that the precipitation response to forest loss was robustly negative during both El Niño and non-El Niño years (Extended Data Fig. 3). Over the Amazon and SEA, we see a stronger reduction in precipitation over regions of forest loss during El Niño years. The relative impact of El Niño on precipitation is smaller in the Congo[22], and correspondingly we do not see a stronger reduction here. A stronger precipitation response to forest loss in regions and periods impacted by El Niño is probably due to higher transpiration rates observed in tropical forests during El Niño years[23] and because rainfall is more sensitive to reductions in moisture recycling during drought years[5,24]. Climate change is expected to lead to increased droughts over many tropical regions[25], which may be further exacerbated by ongoing deforestation.

Seasonal precipitation reductions

Changes in precipitation due to forest loss during the dry, wet and transition seasons are nearly consistently negative (Fig. 3). For the tropics, absolute changes in precipitation with forest loss are greatest in the wet season (Fig. 3a, up to −0.6 mm per month per percentage point forest loss) whereas

relative changes of precipitation with forest loss are similar (−0.2% per percentage point) across dry, wet and transition seasons (Supplementary Fig. 2). In the Amazon, deforestation causes the largest reductions in precipitation during the transition season (Fig. 3b) as has been found previously[18,26,27].

Fig. 3: Changes in seasonal precipitation due to forest loss.



a–**d**, Bars indicate the median change in precipitation (millimetres per month) per percentage point forest cover loss for satellite datasets during 2003–2017 for tropics (**a**), Amazon (**b**), Congo (**c**) and SEA (**d**). Error bars indicate ±1 standard error from the mean. Statistically significant (*$P < 0.05$; **$P < 0.01$) and nonsignificant (NS) differences in changes in mean

precipitation over deforested regions compared with controls are indicated. Results are shown for the wettest 3 months (wet), the driest 3 months (dry) and the transition months (remaining 6 months). Datasets used in this analysis detailed in Extended Data Table 1.

Source Data

Previous case studies have indicated that dry-season precipitation can increase over deforestation in the Amazon[11,28,29]. We observed a nonsignificant increase in dry-season precipitation due to forest loss in the Amazon at 2° as well as increases in the Congo at 1° and 2° (Fig. 3). In SEA, forest loss causes reductions in dry-season precipitation across all scales (Fig. 3d). The mechanism through which forest loss impacts precipitation is likely to change with both season and spatial scale. At the smallest scales (5 km), thermally driven impacts are likely to dominate, shifting to dynamically driven impacts through reductions to surface roughness, then to reductions in moisture fluxes and precipitation recycling at the largest scales[12,30]. Our observation of greater reductions in precipitation due to deforestation at larger spatial scales is consistent with a reduction in moisture recycling emerging as the dominant mechanism[1].

Comparison with climate models

A meta-analysis of climate model studies (predominantly global models with >2° resolution) found that forest loss in the Amazon resulted in a mean reduction in annual mean precipitation of 0.16 ± 0.13% per percentage point[17], overlapping with our value of 0.25% per percentage point (Supplementary Fig. 2). Fewer simulations have been conducted for the Congo, with models predicting a reduction in precipitation of 0.16 ± 0.17% per percentage point[2], similar to our reduction of 0.15% per percentage point (Supplementary Fig. 2). The large range of model estimates highlights the substantial uncertainty in model predictions. Our observationally derived analysis provides support for models that predict reductions in precipitation under regional deforestation at global climate model scales.

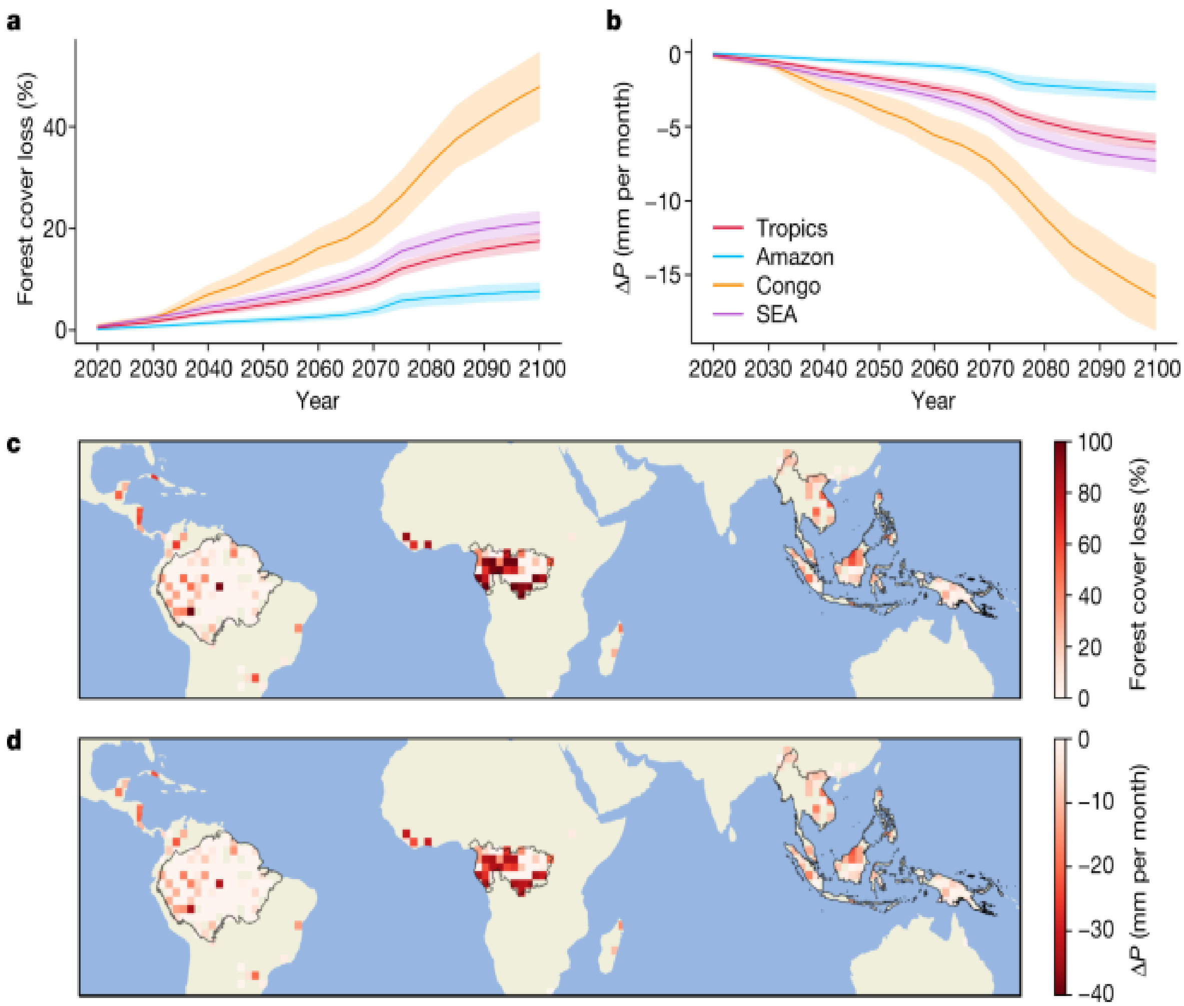
Our observational analysis documents the impacts of deforestation on precipitation across the tropics. Applying linear scaling to the reductions in

precipitation observed in our analysis would suggest that complete deforestation could result in reductions in annual precipitation of 10–20%. Previous estimates of the impact of complete deforestation on precipitation range from a 16% (ref. [17]) to 55–70% (ref. [31]) reduction in the Amazon and an 18% (ref. [2]) to 50% (ref. [32]) reduction in the Congo.

Impacts of future deforestation

To further explore how future deforestation might modify precipitation, we combined our observationally derived estimates of precipitation responses to forest cover loss with future projections of land cover change from a high-deforestation scenario (Methods). We estimate that forest loss from 2015 to 2100 (Fig. 4a) could lead to reductions of annual mean precipitation of up to 16.5 ± 6.2 mm per month in the Congo (Fig. 4b), equivalent to precipitation declines of 8–10%. Forest loss is projected to be greatest in the western and southern Congo (Fig. 4c), which will also experience the strongest reductions in precipitation (Fig. 4d).

Fig. 4: Impact of projected future forest loss on annual mean precipitation.



a, Mean forest cover loss over 2015–2100 under Shared Socioeconomic Pathway 3–Representative Concentration Pathway 4.5 for the tropics, Amazon, Congo and SEA. **b**, Impact of projected forest cover loss on precipitation (*P*; ±1 standard error from the mean). **c**, Spatial pattern of forest cover loss. **d**, Predicted *P* change (Δ*P*) in 2100 due to forest cover loss. Results are shown for 2.0° resolution. Maps of the different regions generated using Cartopy and Natural Earth[51].

Source Data

The sensitivity of precipitation to the extent of forest loss is an uncertainty in our analysis, a result of the relatively short observational record, compounded by large spatial and temporal variability in precipitation. The response of precipitation to forest loss greater than 30%, a threshold beyond which large reductions in precipitation have been postulated[1], is one such

uncertainty. Restricting our analysis to the forest losses of 0–30% that are well sampled in our observational dataset (Supplementary Fig. 3), through capping the impacts of greater forest loss at that of 30%, results in projected annual mean precipitation reductions of 6.5 ± 2.6 mm per month in the Congo and 6.2 ± 2.5 mm per month in SEA (Supplementary Fig. 4). However, restricting our analysis in this way is likely to underestimate the precipitation impacts over regions projected to experience the most extensive deforestation, including the Congo, where mean forest cover is projected to decline by 40 percentage points between 2015 and 2100 (Fig. 4a).

Previous studies have identified both linear[9,33] and nonlinear[1,31] responses of precipitation to forest loss. Such nonlinear interactions and feedbacks have the potential to further amplify or moderate the responses predicted here[14,34]. Our analysis shows large reductions in precipitation for relatively small amounts of forest loss and evidence for reduced sensitivity of precipitation to additional amounts of forest loss (Extended Data Fig. 1). Assuming a nonlinear relationship between forest loss and precipitation (Methods) reduces our projected reductions in precipitation by around a factor of 2 (Supplementary Fig. 5). Our observationally based approach will miss tipping points in the climate system that might be reached as deforestation extent progresses further[1]. Such tipping points have been postulated for the Amazon under future global change[25,35]. Thus, the substantial declines in precipitation projected in our analysis should be viewed as a conservative estimate of the potential precipitation response to future deforestation. Nevertheless, our analysis suggests that deforestation can drive local and regional precipitation changes that may match or exceed those predicted due to climate change over the same period[36,37].

Implications of precipitation reductions

Reductions in precipitation induced by forest loss have important implications for society and the sustainability of remaining tropical forest. Deforestation-induced reductions in precipitation affect agriculture[1,14] and hydropower generation[38]. On average, crop yields decline by 0.5% for each percentage point reduction in precipitation[39]. Our results indicate that forest-

loss-induced changes to annual precipitation (Supplementary Fig. 2) could cause crop yields to decline by 1.25% for each 10-percentage-point loss of forest cover, potentially exacerbating the impacts of climate change and future drought events. The maintenance of regional rainfall patterns due to forests in the Amazon has been valued at up to US$9 ha^{-1} yr^{-1} and US$1.84 ha^{-1} yr^{-1} through sustaining agricultural yields and hydropower generation, respectively[40]. Global cropland area increased by 9% in the past two decades, with even higher increases in South America and tropical Africa[41] largely at the expense of natural ecosystems. Further agricultural expansion in tropical forest regions may lead to overall reductions in production if declines in yield due to deforestation-induced reductions in rainfall outweigh increased production from expanded agricultural area[14].

Furthermore, reductions in rainfall over remaining areas of tropical forest are expected to lead to additional forest loss[9] as well as impacting species composition[22], carbon sequestration[42] and fire frequency[43]. Reductions in dry-season precipitation pose a particular threat to forest viability by exacerbating seasonal droughts and potentially delaying the onset of the wet season and extending the length of the dry season. Increases in dry-season length over recent decades have previously been reported for the Amazon[44] and the Congo[45], possibly linked to land cover changes[27].

Deforestation may also shift precipitation patterns, increasing dry-season rainfall immediately downwind of forest loss and decreasing rainfall in upwind areas[12]. Our approach is restricted to observing deforestation impacts up to scales of 200 km (Methods). At larger scales, insufficient pixels experienced forest loss during the relatively short period of satellite observations for a robust analysis. Deforestation is also likely to alter precipitation at these larger scales through reducing moisture recycling leading to reductions in rainfall downwind of forest loss[4,5,9,35]. The length scale of moisture recycling has been estimated at 500–2,000 km in the tropics[46], with a median value of 600 km in the Amazon[5]. In regions downwind of extensive forests, such as the southwestern Amazon, up to 70% of precipitation could be sourced from upwind evapotranspiration[47,48]. Tropical forest loss could therefore have severe implications for precipitation in these regions that are hundreds to thousands of kilometres

downwind of the forest loss[5]. Through missing the impacts at these larger scales, our analysis is likely to underestimate the full impacts of deforestation on rainfall.

Our results highlight the importance of remaining tropical forests for sustaining regional precipitation. Despite efforts to reduce deforestation, rates of tropical forest loss have accelerated over the past two decades[49]. Renewed efforts are needed to ensure recent commitments to reduce deforestation, including the New York Declaration on Forests and The Glasgow Leaders' Declaration on Forests and Land Use made at the 26th UN Climate Change Conference of the Parties, are successful. Global efforts to restore large areas of degraded and deforested land could enhance precipitation[50], reversing some of the reductions in precipitation due to forest loss observed here.

Methods

Datasets

We used 18 precipitation datasets, listed in Extended Data Table 1. All datasets were downloaded at the highest available spatial resolution, which for some datasets was 0.04°, or approximately 4 km at the Equator. Data were obtained as monthly means or converted to monthly mean using the Python package xarray[52]. We categorized precipitation datasets as satellite ($n = 10$), station ($n = 4$) and reanalysis ($n = 4$). Satellite datasets are those based primarily on data from satellite sensors and include datasets that have both satellite and station-based data (that is, merged datasets). Station datasets include only ground-based information from weather stations and rain gauges. Reanalysis products are models constrained by surface and satellite data. Precipitation datasets have been compared previously over the Amazon[20] highlighting the limited station data over tropical forest regions. Time series of precipitation (Supplementary Fig. 1) reveal variability across the different datasets highlighting the need to analyse impacts of deforestation across multiple datasets.

To analyse the changes in forest canopy cover, we used data from the Global Forest Change (GFC) version 1.9 (ref. [10]). GFC v1.9 provides forest canopy

cover in the year 2000 and subsequent annual forest loss from 2001–2020 at 30-m resolution. We analysed forest cover and precipitation changes over the period 2003 to 2017, which was the period common to all datasets.

Analysis across multiple spatial scales

We analysed the impacts of forest loss across a range of scales (0.05°, 0.1°, 0.25°, 0.5°, 1.0° and 2.0°). Each precipitation dataset was analysed at its native resolution and at all lower resolutions across this range of scales. Spatial regridding was carried out using the Python package xESMF[53] with a bilinear regridding scheme. Two alternative regridding methods (xESMF: conservative-normalized; and iris: area weighted) were tested and had little impact on our results. For GFC data, we calculated forest loss using the original 30-m data and converted the resulting values to each of the six spatial resolutions analysed by taking the sum of all 30-m pixels within each larger pixel. Change in canopy cover from 2003 to 2017 at each resolution is shown in Fig. 1.

Assessing impact of historical deforestation on precipitation

We used a moving-window nearest-neighbour approach[54] to compare the forest loss and precipitation change of each pixel with that of its immediate neighbours. We tested the sensitivity of the analysis to the size of the moving window and found similar results for 3 × 3 and 5 × 5 (Extended Data Fig. 2) moving windows. Results from the 3 × 3 moving-window approach can been seen in the main paper. We calculated the forest loss of each deforested pixel relative to neighbouring control pixels as the forest loss of the deforested pixel minus forest loss of the control. We constrained our analysis to the tropical evergreen broadleaf biome using the Moderate Resolution Imaging Spectroradiometer land cover dataset[55]. To be included in the analysis, deforested pixels must have experienced 0.1% more forest loss over time than their neighbouring control pixels. The number of deforested pixels analysed varied between analysis resolutions as follows: 0.05°, $n = 243{,}254$; 0.1°, $n = 58{,}660$; 0.25°, $n = 9{,}604$; 0.5°, $n = 2{,}303$; 1.0°, $n = 586$; 2.0°, $n = 123$. We observed similar distributions of canopy change for all spatial resolutions analysed (Supplementary Fig. 6).

We calculated the precipitation change of the deforested pixel relative to the precipitation change of the control pixel (ΔP) as the precipitation change of the deforested pixel over the analysis period (for example, 2003–2017) minus the precipitation change over the control pixel. To reduce the impact of interannual variability in precipitation on our results, we calculated 5-yr means for periods at the start (2003–2007) and end (2013–2017; Extended Data Fig. 5) of the analysis period. We then calculated the change in precipitation as the difference between the start and end of these multi-year means. We report precipitation changes (ΔP) as a function of forest loss by dividing by the difference in forest loss between deforestation and control pixels (units of millimetres per month per percentage point). We also report precipitation change as the percentage change in precipitation ($\Delta P/P$, in units of per cent) as a function of forest loss (in units of per cent per percentage point).

To ensure that control pixels and deforested pixels experience a similar background climate, we conducted a sensitivity test in which we restricted our analysis to pixels for which the pre-deforestation precipitation across the control and deforested pixels differed by less than 10%. Restricting our analysis in this way had little impact on our results (Supplementary Fig. 7) showing that our nearest-neighbour approach is effective even at the largest scales analysed here.

To explore the role of the analysis period on our results, we compared the results for 5-yr means to those for shorter 3-yr means (2003–2005 versus 2015–2017) and found consistent results (Extended Data Fig. 3). Our analysis period includes the strong 2015/2016 El Niño that resulted in reductions in precipitation over most tropical land regions, particularly in 2015 (Supplementary Fig. 1). To explore the potential impacts of the 2015/2016 El Niño on our analysis, we estimated the impact of forest loss on precipitation using 3-yr (2003–2005 versus 2018–2020) and 5-yr (2003–2007 versus 2016–2020) multi-annual means spanning an extended time period. The 3-yr analysis completely excludes the 2015/2016 ENSO, and the 5-yr analysis excludes 2015, which was the driest year (Extended Data Fig. 3). Two datasets (TRMM and UDEL) were not available after 2017, so they were removed from this sensitivity analysis.

Statistical analysis

For each category of precipitation data (satellite, station and reanalysis), precipitation change values were grouped together for all deforestation pixels and all control pixels. We found that precipitation changes for deforested pixels and control pixels, and the difference in precipitation change between deforested and control pixels (Extended Data Fig. 4), were normally distributed. Error bars (Figs. 2 and 3) show ±1 standard error from the mean calculated and displayed using the Python package Seaborn[56]. To test whether mean precipitation changes over regions of deforestation were statistically different from changes over the control areas, we used a Student's *t*-test. We also used the Mann–Whitney test to test for significant differences in median precipitation change between control and deforested pixels and found similar results.

Seasonal analysis

For the satellite datasets alone, in addition to calculating precipitation changes at the annual timescale, we calculated changes for the dry season (driest 3 months of each year), wet season (wettest 3 months of each year) and transition season (remaining 6 months). The driest, wettest and transition months were identified for each pixel using each individual precipitation dataset. For each season and dataset, we calculated the median change in precipitation across all of the pixels within the region of interest (Supplementary Figs. 8–10).

Predicting future precipitation change due to forest loss

We used projections of forest cover change available at 0.05° from the Global Change Analysis Model (GCAM) for 2015–2100 based on the Shared Socioeconomic Pathway 3–Representative Concentration Pathway 4.5 scenario, which represents a high-deforestation future[57]. GCAM includes the impacts of climate and land use on future forest cover. We summed forest cover from all forest categories and calculated forest cover loss in each year compared to a 2015 baseline. Forest cover loss data were regridded to 2°. We estimated the impact of forest loss on future precipitation at the 2° scale through multiplying the projected percentage point forest loss for each pixel by the observed median change in precipitation (millimetres per month) per percentage point forest cover loss

across the satellite datasets. To estimate the uncertainty in our predictions, we applied an upper and lower limit on the sensitivity of precipitation to forest loss based on the median value ±1 standard error from the mean (see error bars in Fig. 2) and rescaled by forest loss. This provides a range of estimated precipitation impacts of future forest loss. We also tested the impact on our results of capping future forest loss in each pixel at 30%, which is the upper range of forest loss that is well sampled in the observations (Supplementary Fig. 3). For each region, we applied the tropical satellite precipitation response to forest loss (Fig. 2f), meaning that our projected regional precipitation changes are a product of the regional canopy cover change and the median tropical precipitation response. Our approach assumes a linear precipitation response to forest loss, which recent work suggests could provide a conservative estimate of deforestation impacts[31]. We tested the sensitivity of assuming a linear response of precipitation to canopy cover loss. We fitted a nonlinear function to the data presented in Extended Data Fig. 1 through applying the median sensitivity of precipitation to forest cover loss (millimetres per month per percentage point) within each forest cover loss bin. We then scaled by the projected forest cover loss. This approach reduces the projected reduction in precipitation to 2.4 mm per month in SEA and 1.5 mm per month in the Congo (Supplementary Fig. 5).

Data availability

Full results for all tested resolutions used in this analysis are available through https://doi.org/10.5281/zenodo.7373832. The original datasets are freely available to download from the following repositories: CHIRPS from https://data.chc.ucsb.edu/products/?C=M;O=D, CMORPH from https://ftp.cpc.ncep.noaa.gov/precip/CMORPH_RT/GLOBE/data/, CPC from https://psl.noaa.gov/data/gridded/data.cpc.globalprecip.html, CRU from https://crudata.uea.ac.uk/cru/data/hrg/, ERA5 from https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview, GPCC from https://opendata.dwd.de/climate_environment/GPCC/html/download_gate.html, GPCP from https://disc.gsfc.nasa.gov/datasets/GPCPMON_3.1/summary?keywords=GPCPMON, GPM from

https://gpm1.gesdisc.eosdis.nasa.gov/data/GPM_L3/, JRA from https://climatedataguide.ucar.edu/climate-data/jra-55 and https://jra.kishou.go.jp/JRA-55/index_en.html, MERRA-2 from https://disc.gsfc.nasa.gov/datasets?project=MERRA-2, NOAA (PREC/LAND) from https://psl.noaa.gov/data/gridded/data.precl.html, PERSIANN (CCS, CDR, CCS-CDR, PDIR-NOW) from https://chrsdata.eng.uci.edu/, TRMM from https://disc.gsfc.nasa.gov/datasets/TRMM_3B43_7/summary, UDEL from https://psl.noaa.gov/data/gridded/data.UDel_AirT_Precip.html. The GCAM model output used in this study is available from https://doi.org/10.25584/data.2020-07.1357/1644253. Source data are provided with this paper.

Code availability

The code used in this analysis is available through https://doi.org/10.5281/zenodo.7373832.

References

1. Lawrence, D. & Vandecar, K. Effects of tropical deforestation on climate and agriculture. *Nat. Clim. Change* **5**, 27–36 (2015).

2. Spracklen, D. V., Baker, J. C. A., Garcia-Carreras, L. & Marsham, J. H. The effects of tropical vegetation on rainfall. *Annu. Rev. Environ. Resour.* **43**, 193–218 (2018).

3. Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008).

4. Spracklen, D. V., Arnold, S. R. & Taylor, C. M. Observations of increased tropical rainfall preceded by air passage over forests. *Nature* **489**, 282–285 (2012).

5. Staal, A. et al. Forest-rainfall cascades buffer against drought across the Amazon. *Nat. Clim. Change* **8**, 539–543 (2018).

6. Baker, J. C. A. & Spracklen, D. V. Divergent representation of precipitation recycling in the Amazon and the Congo in CMIP6 models. *Geophys. Res. Lett.* **49**, e2021GL095136 (2022).

7. Guan, K. et al. Photosynthetic seasonality of global tropical forests constrained by hydroclimate. *Nat. Geosci.* **8**, 284–289 (2015).

8. Staal, A. et al. Hysteresis of tropical forests in the 21st century. *Nat. Commun.* **11**, 4978 (2020).

9. Zemp, D. C. et al. Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks. *Nat. Commun.* **8**, 14681 (2017).

10. Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–854 (2013).

11. Chagnon, F. J. F. & Bras, R. L. Contemporary climate change in the Amazon. *Geophys. Res. Lett.* **32**, L13703 (2005).

12. Khanna, J., Medvigy, D., Fueglistaler, S. & Walko, R. Regional dry-season climate changes due to three decades of Amazonian deforestation. *Nat. Clim. Change* **7**, 200–204 (2017).

13. Garcia-Carreras, L. & Parker, D. J. How does local tropical deforestation affect rainfall? *Geophys. Res. Lett.* **38**, L19802 (2011).

14. Leite-Filho, A. T., Soares-Filho, B. S., Davis, J. L., Abrahão, G. M. & Börner, J. Deforestation reduces rainfall and agricultural revenues in the Brazilian Amazon. *Nat. Commun.* **12**, 2591 (2021).

15. McAlpine, C. A. et al. Forest loss and Borneo's climate. *Environ. Res. Lett.* **13**, 044009 (2018).

16. Chapman, S. et al. Compounding impact of deforestation on Borneo's climate during El Niño events. *Environ. Res. Lett.* **15**, 084006 (2020).

17. Spracklen, D. V. & Garcia-Carreras, L. The impact of Amazonian deforestation on Amazon basin rainfall. *Geophys. Res. Lett.* **42**, 9546–9552 (2015).

18. Jiang, Y. et al. Modeled response of South American climate to three decades of deforestation. *J. Clim.* **34**, 2189–2203 (2021).

19. Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* **7**, 109 (2020).

20. Fassoni-Andrade, A. C. et al. Amazon hydrology from space: scientific advances and future challenges. *Rev. Geophys.* **59**, e2020RG000728 (2021).

21. Haiden, T., Janousek, M., Vitart, F., Ferranti, L. & Prates, F. *Evaluation of ECMWF Forecasts, Including the 2019 Upgrade*. *ECMWF Technical Memorandum No. 853* (ECMWF, 2019).

22. Esquivel-Muelbert, A. et al. Compositional response of Amazon forests to climate change. *Glob. Change Biol.* **25**, 39–56 (2019).

23. Brum, M. et al. ENSO effects on the transpiration of eastern Amazon trees. *Philos. Trans. R. Soc. B* **373**, 20180085 (2018).

24. Bagley, J. E., Desai, A. R., Harding, K. J., Snyder, P. K. & Foley, J. A. Drought and deforestation: has land cover change influenced recent precipitation extremes in the Amazon? *J. Clim.* **27**, 345–361 (2014).

25. Wunderling, N. et al. Recurrent droughts increase risk of cascading tipping events by outpacing adaptive capacities in the Amazon rainforest. *Proc. Natl Acad. Sci. USA* **119**, e2120777119 (2022).

26. Fu, R. & Li, W. The influence of the land surface on the transition from dry to wet season in Amazonia. *Theor. Appl. Climatol.* **78**, 97–110 (2004).

27. Leite-Filho, A. T., de Sousa Pontes, V. Y. & Costa, M. H. Effects of deforestation on the onset of the rainy season and the duration of dry spells in southern Amazonia. *J. Geophys. Res. Atmos.* **124**, 5268–5281 (2019).

28. Negri, A. J., Adler, R. F., Xu, L. & Surratt, J. The Impact of Amazonian deforestation on dry season rainfall. *J. Clim.* **17**, 1306–1319 (2004).

29. Chagnon, F. J. F., Bras, R. L. & Wang, J. Climatic shift in patterns of shallow clouds over the Amazon. *Geophys. Res. Lett.* **31**, L24212 (2004).

30. Chambers, J. Q. & Artaxo, P. Biosphere–atmosphere interactions: deforestation size influences rainfall. *Nat. Clim. Change* **7**, 175–176 (2017).

31. Baudena, M., Tuinenburg, O. A., Ferdinand, P. A. & Staal, A. Effects of land-use change in the Amazon on precipitation are likely underestimated. *Glob. Change Biol.* **27**, 5580–5587 (2021).

32. Duku, C. & Hein, L. The impact of deforestation on rainfall in Africa: a data-driven assessment. *Environ. Res. Lett.* **16**, 064044 (2021).

33. Akkermans, T., Thiery, W. & Van Lipzig, N. P. M. The regional climate impact of a realistic future deforestation scenario in the Congo basin. *J. Clim.* **27**, 2714–2734 (2014).

34. Staal, A. et al. Feedback between drought and deforestation in the Amazon. *Environ. Res. Lett.* **15**, 044024 (2020).

35. Xu, X. et al. Deforestation triggering irreversible transition in Amazon hydrological cycle. *Environ. Res. Lett.* **17**, 034037 (2022).

36. Kooperman, G. J. et al. Forest response to rising CO_2 drives zonally asymmetric rainfall change over tropical land. *Nat. Clim. Change* **8**, 434–440 (2018).

37. Chen, Z. et al. Global land monsoon precipitation changes in CMIP6 projections. *Geophys. Res. Lett.* **47**, e2019GL086902 (2020).

38. Stickler, C. M. et al. Dependence of hydropower energy generation on forests in the Amazon Basin at local and regional scales. *Proc. Natl Acad. Sci. USA* **110**, 9601–9606 (2013).

39. Challinor, A. J. et al. A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Change* **4**, 287–291 (2014).

40. Strand, J. et al. Spatially explicit valuation of the Brazilian Amazon forest's ecosystem services. *Nat. Sustain.* **1**, 657–664 (2018).

41. Potapov, P. et al. Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nat. Food* **3**, 19–28 (2022).

42. Li, Y. et al. Deforestation-induced climate change reduces carbon storage in remaining tropical forests. *Nat. Commun.* **13**, 1964 (2022).

43. Aragão, L. E. O. C. et al. Interactions between rainfall, deforestation and fires during recent years in the Brazilian Amazonia. *Philos. Trans. R. Soc. B* **363**, 1779–1785 (2008).

44. Marengo, J. A. et al. Changes in climate and land use over the Amazon region: current and future variability and trends. *Front. Earth Sci.* https://doi.org/10.3389/feart.2018.00228 (2018).

45. Jiang, Y. et al. Widespread increase of boreal summer dry season length over the Congo rainforest. *Nat. Clim. Change* https://doi.org/10.1038/s41558-019-0512-y (2019).

46. Van Der Ent, R. J. & Savenije, H. H. G. Length and time scales of atmospheric moisture recycling. *Atmos. Chem. Phys.* **11**, 1853–1863 (2011).

47. Sorí, R., Nieto, R., Vicente-Serrano, S. M., Drumond, A. & Gimeno, L. A Lagrangian perspective of the hydrological cycle in the Congo River basin. *Earth Syst. Dyn.* **8**, 653–675 (2017).

48. van der Ent, R. J., Savenije, H. H. G., Schaefli, B. & Steele-Dunne, S. C. Origin and fate of atmospheric moisture over continents. *Water Resour. Res.* **46**, W09525 (2010).

49. Feng, Y. et al. Doubling of annual forest carbon loss over the tropics during the early twenty-first century. *Nat. Sustain.* **4**, 441–451 (2022).

50. Tuinenburg, O. A., Bosmans, J. H. C. & Staal, A. The global potential of forest restoration for drought mitigation. *Environ. Res. Lett.* **17**, 034045 (2022).

51. Met Office. Cartopy: a cartographic python library with a Matplotlib interface 2010–2015. *Met Office* https://scitools.org.uk/cartopy (2022).

52. Hoyer, S. & Hamman, J. xarray: N-D labeled arrays and datasets in Python. *J. Open Res. Softw.* https://doi.org/10.5334/jors.148 (2017).

53. Zhuang, J. xESMF. *Zenodo* https://doi.org/10.5281/zenodo.1134365 (2022).

54. Baker, J. C. A. & Spracklen, D. V. Climate benefits of intact Amazon forests and the biophysical consequences of disturbance. *Front. For. Glob. Change* https://doi.org/10.3389/ffgc.2019.00047 (2019).

55. Schaaf, C. & Wang, Z. MCD43A3 MODIS/Terra+Aqua BRDF/Albedo Daily L3 Global - 500m V006. *NASA EOSDIS Land Processes DAAC* https://doi.org/10.5067/modis/mcd43a3.006 (2015).

56. Waskom, M. Seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).

57. Chen, M. et al. Global land use for 2015–2100 at 0.05° resolution under diverse socioeconomic and climate scenarios. *Sci. Data* **7**, 320 (2020).

58. Funk, C. et al. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci. Data* **2**, 150066 (2015).

59. Xie, P. et al. NOAA Climate Data Record (CDR) of CPC Morphing technique (CMORPH) high resolution global precipitation estimates, version 1. *NOAA National Centers for Environmental Information* https://doi.org/10.25921/w9va-q159 (2019).

60. Xie, P. et al. A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeorol.* **8**, 607–626 (2007).

61. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).

62. Elke, R., Hänsel, S., Finger, P., Schneider, U. & Ziese, M. GPCC Climatology Version 2022 at 0.25°: monthly land-surface precipitation climatology for every month and the total year from rain-gauges built on GTS-based and historical data. *GPCC* https://doi.org/10.5676/DWD_GPCC/CLIM_M_V2022_025 (2022).

63. Huffman, G. J. A., Behrangi, R. F., Adler, D. T., Bolvin, E. J. & Nelkin, G. G. Introduction to the new version 3 GPCP monthly global precipitation analysis. *GPCP* https://docserver.gesdisc.eosdis.nasa.gov/public/project/MEaSUREs/GPCP/Release_Notes.GPCPV3.2.pdf (2022).

64. Hou, A. Y. et al. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **95**, 701–722 (2014).

65. Kobayashi, S. et al. The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteorol. Soc. Japan* **93**, 5–48 (2015).

66. Gelaro, R. et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).

67. Chen, M., Xie, P. & Janowiak, J. E. Global land precipitation: a 50-yr monthly analysis based on gauge observations. *J. Hydrometeorol.* **3**, 249–266 (2002).

68. Nguyen, P. et al. The CHRS data portal, an easily accessible public repository for PERSIANN global satellite precipitation data. *Sci. Data* **6**, 1180296 (2019).

69. Ashouri, H. et al. PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **96**, 69–83 (2015).

70. Nguyen, P. et al. Persiann dynamic infrared–rain rate (PDIR-now): a near-real-time, quasi-global satellite precipitation dataset. *J. Hydrometeorol.* **21**, 2893–2906 (2020).

71. Sadeghi, M. et al. PERSIANN-CCS-CDR, a 3-hourly 0.04° global precipitation climate data record for heavy precipitation studies. *Sci. Data* **8**, 157 (2021).

72. Huffman, G. J. et al. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8**, 38–55 (2007).

73. Matsuura, K. & Willmott, C. J. Terrestrial precipitation: 1900-2017 gridded monthly time series. *Global Precipitation Archive* http://climate.geog.udel.edu/~climate/html_pages/Global2017/README.GlobalTsP2017.html (2018).

Acknowledgements

The research has been supported by funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (DECAF project, grant agreement no. 771492), and the Newton Fund, through the Met Office Climate Science for Service Partnership Brazil.

Author information

Authors and Affiliations

1. School of Earth and Environment, University of Leeds, Leeds, UK

   C. Smith, J. C. A. Baker & D. V. Spracklen

Contributions

All authors developed the concept of the study, and contributed to experimental design, interpretation of results and writing of the manuscript. C.S. and J.C.A.B. contributed to the analysis.

Corresponding author

Correspondence to C. Smith.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Arie Staal and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Annual precipitation change as a function of forest loss.

Results are shown at 2° spatial resolution for all satellite precipitation (P) datasets calculated as the change in P over time for deforested data pixels minus change over time for control data pixels. Data is binned according to forest cover change (%) with an equal number of pixels in each bin. Points show the median and error bars show ± 1 standard error from the mean. Details of each data product are provided in Extended Data Table 1.

Source Data

Extended Data Fig. 2 Annual precipitation change due to forest loss for individual datasets.

Results are shown for 2003 – 2017 for 5 year averages and 3x3 moving window. Bars show the median absolute change in annual P (mm month^{-1}) per percentage point tree cover loss in each region (Tropics (a-f), Amazon (g-l), Congo (m-r), SEA (s-x)). Each P dataset is shown separately and ordered and coloured by category: satellite (orange), station (yellow) and reanalysis (turquoise). The datasets are numbered; 1) CHIRPS, 2) CMORPH, 3) CPC, 4) CRU, 5) ERA5, 6) GPCC, 7) GPCP, 8) GPM, 9) JRA, 10) MERRA-2, 11) NOAA 12) PERSIANN-CCS, 13) PERSIANN-CCSCDR, 14) PERSIANN-CDR, 15) PERSIANN-NOW, 16) PERSIANN, 17) TRMM, 18) UDEL. Results are shown for forest loss scales of 0.05° (a,g,m,s), 0.1° (b,h,n,t), 0.25° (c,i,o,u), 0.5° (d,j,p,v), 1.0° (e,k,q,w), 2.0° (f,l,r,x). Details of each data product are provided in Extended Data Table 1.

Source Data

Extended Data Fig. 3 Changes in precipitation due to forest loss for different time periods and nearest neighbour comparisons.

Changes in annual mean precipitation at 2.0° resolution are shown for satellite (orange), station (yellow) and reanalysis (turquoise) datasets for the tropics (a-f), Amazon (g-l), Congo (m-r) and Southeast Asia (SEA, s-x). Columns show the sensitivity of our results to changes in the analysis period, number of years used to compute multi-annual means at start and end of the analysis period, and size of the moving window used for nearest neighbour comparisons: 2003-2017, 3-year averages and 3x3 nearest neighbour (Column 1, a,g,m,s); 2003-2017, 3-year, 5x5 (Column 2; b,g,n,t); 2003-2017, 5-year, 3x3 (Column 3; c,i,o,u); 2003-2017, 5-year, 5x5 (Column 4; d,j,p,v); 2003-2020, 3-year, 3x3 (Column 5; e,k,q,w); 2003-2020, 5-year, 3x3 (Column 6; f,l,r,x). Error bars show ± 1 standard error from the mean. Details of each data product are provided in Extended Data Table 1. Full results for all tested resolutions are available in an online repository [10.5281/zenodo.7373832].

Source Data

Extended Data Fig. 4 Change in precipitation over deforested, control and difference between deforested and control pixels.

Change in precipitation over 2003 to 2017 is shown for deforested (a, b), control (c, d) and difference between deforested and control pixels (e, f) for 0.05° (a, c, e) and 2.0° (b, d, f) resolution. Details of each data product are provided in Extended Data Table 1.

Source Data

Extended Data Fig. 5 Mean precipitation from satellite, station and reanalysis datasets.

For each class of dataset, satellite (a, d, g), station (b, e, h) and reanalysis (c, f, i), the median value for the 5-year multi-annual mean at the start (2003-2007; a, b, c) and end (2013-2017; d, e, f) of the analysis period as well as the change over the analysis period (end – start; g, h, i) is shown. Mean values across tropical evergreen broadleaf forests are shown in units of mm/month at the top of each panel. Maps of the different regions generated using Cartopy and Natural Earth[51]. Details of each data product are provided in Extended Data Table 1.

Source Data

Extended Data Table 1 Precipitation datasets used in this study refs [58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73]

Supplementary information

Supplementary Figures

This file Supplementary Figs. 1–10, which support the results presented in the main paper. Here we also present results from the individual precipitation datasets.

Peer Review File

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from

the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Smith, C., Baker, J.C.A. & Spracklen, D.V. Tropical deforestation causes large reductions in observed precipitation. *Nature* **615**, 270–275 (2023). https://doi.org/10.1038/s41586-022-05690-1

- Received: 14 April 2022
- Accepted: 15 December 2022
- Published: 01 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-022-05690-1

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

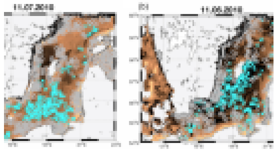
Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Observed reductions in rainfall due to tropical deforestation

Nature Research Briefing 01 Mar 2023

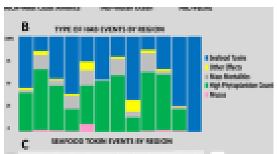
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-022-05690-1



Retracing cyanobacteria blooms in the Baltic Sea

27 June 2022

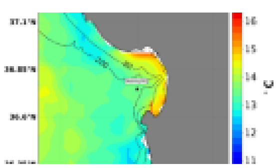
U. Löptien & H. Dietze



Perceived global increase in algal blooms is attributable to intensified monitoring and emerging bloom impacts

08 June 2021

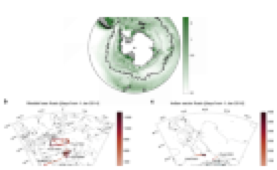
Gustaaf M. Hallegraeff, Donald M. Anderson, … Adriana Zingone



Downscaling global ocean climate models improves estimates of exposure regimes in coastal environments

26 August 2020

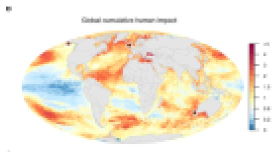
Matheus Fagundes, S. Y. Litvin, … C. B. Woodson



[Remote assessment of the fate of phytoplankton in the Southern Ocean sea-ice zone]

19 June 2020

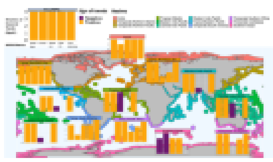
Sébastien Moreau, Philip W. Boyd & Peter G. Strutton



[Recent pace of change in human impact on the world's ocean]

12 August 2019

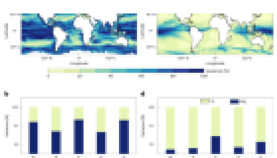
Benjamin S. Halpern, Melanie Frazier, … Kimberly A. Selkoe



[Unravelling seasonal trends in coastal marine heatwave metrics across global biogeographical realms]

11 May 2022

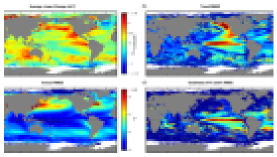
François Thoral, Shinae Montie, … David R. Schiel



[Annual variations in phytoplankton biomass driven by small-scale physical processes]

14 November 2022

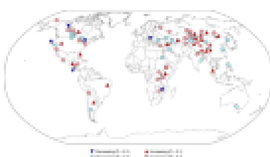
M. G. Keerthi, C. J. Prend, … M. Lévy



[Global patterns of change and variation in sea surface temperature and chlorophyll a]

02 October 2018

Piers K. Dunstan, Scott D. Foster, … Peter A. Thompson



[Widespread global increase in intense lake phytoplankton blooms since the 1980s]

14 October 2019

Jeff C. Ho, Anna M. Michalak & Nima Pahlevan

- Article
- Open Access
- Published: 01 March 2023

Coastal phytoplankton blooms expand and intensify in the 21st century

- Yanhui Dai,

- Shangbo Yang,
- Dan Zhao,
- Chuanmin Hu,
- Wang Xu,
- Donald M. Anderson,
- Yun Li,
- Xiao-Peng Song,
- Daniel G. Boyce,
- Luke Gibson,
- Chunmiao Zheng &
- …
- Lian Feng

Show authors

Nature **volume 615**, pages 280–284 (2023)

- 12k Accesses
- 356 Altmetric
- Metrics details

This article has been updated

Abstract

Phytoplankton blooms in coastal oceans can be beneficial to coastal fisheries production and ecosystem function, but can also cause major environmental problems[1,2]—yet detailed characterizations of bloom incidence and distribution are not available worldwide. Here we map daily marine coastal algal blooms between 2003 and 2020 using global satellite observations at 1-km spatial resolution. We found that algal blooms occurred in 126 out of the 153 coastal countries examined. Globally, the spatial extent (+13.2%) and frequency (+59.2%) of blooms increased significantly ($P < 0.05$) over the study period, whereas blooms weakened in tropical and subtropical areas of

the Northern Hemisphere. We documented the relationship between the bloom trends and ocean circulation, and identified the stimulatory effects of recent increases in sea surface temperature. Our compilation of daily mapped coastal phytoplankton blooms provides the basis for global assessments of bloom risks and benefits, and for the formulation or evaluation of management or policy actions.

Main

Phytoplankton blooms are accumulations of microscopic algae in the surface layer of fresh and marine water bodies. Although many blooms can occur naturally, nutrients linked to anthropogenic eutrophication are expected to intensify their frequency globally[2,3,4]. Many algal blooms are beneficial, fixing carbon at the base of the food chain and supporting fisheries and ecosystems worldwide. However, proliferations of algae that cause harm (termed harmful algal blooms (HABs)) have become a major environmental problem worldwide[5,6,7]. For instance, the toxins produced by some algal species can accumulate in the food web, causing closures of fisheries as well as illness or mortality of marine species and humans[8,9,10]. In other cases, the decay of a dense algal bloom can deplete oxygen in bottom waters, forming anoxic 'dead zones' that can cause fish and invertebrate die-offs and ecosystem restructuring, with serious consequences for the well-being of coastal communities[1,11]. Unfortunately, algal bloom frequency and distribution are projected to increase with future climate change[12,13], with some changes causing adverse effects on aquatic ecosystems, fisheries and coastal resources.

Owing to substantial heterogeneity in space and time, algal blooms are challenging to characterize on a large scale[5,14], and thus present knowledge does not allow us to answer one of the most fundamental questions: whether algal blooms have changed in recent decades on a global basis[6,15,16]. For example, although HAB events have been compiled into the UNESCO (United Nations Educational, Scientific, and Cultural Organization) Intergovernmental Oceanographic Commission Harmful Algae Event Database (HAEDAT) globally since 1985, bloom trends are difficult to resolve, owing to inconsistent sampling efforts and the diversity of the eco-

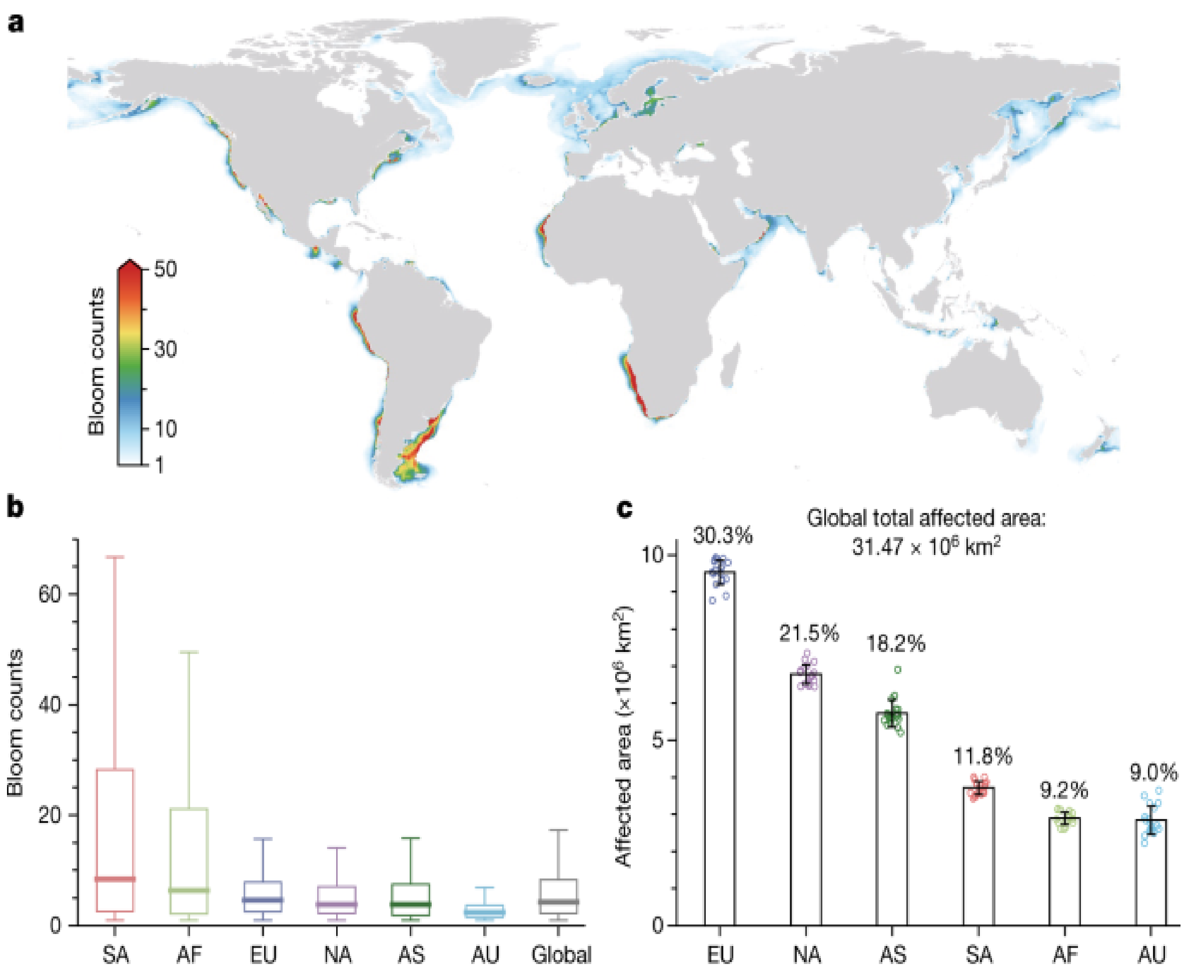
environmental or socio-economic effects[6]. Alternatively, satellite data have been used to monitor the ocean surface continuously since 1997 and have enabled bloom detection in many coastal regions[17,18,19]. However, the technical difficulties in dealing with complex optical features across different types of coastal waters have so far prohibited their application globally[20]. To fill this knowledge gap, we developed a method to map global coastal algal blooms and used this tool to examine satellite images between 2003 and 2020, addressing three fundamental questions: (1) where and how frequently global coastal oceans have been affected by phytoplankton blooms; (2) whether the blooms have expanded or intensified over the past two decades, both globally and regionally; and (3) the identity of the potential drivers.

Mapping global coastal phytoplankton blooms

We generated a satellite-based dataset of phytoplankton bloom occurrence to characterize the spatial and temporal patterns of algal blooms in coastal oceans globally. The dataset was derived using global, 1-km resolution daily observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard NASA's Aqua satellite, and all 0.76 million images acquired by this satellite mission between 2003 and 2020 were used. We developed an automated method to detect phytoplankton blooms using MODIS images (Extended Data Fig. 1) (Methods). In this study, we define a phytoplankton bloom as the accumulation of microscopic algae at the ocean surface that exhibits satellite-detectable fluorescence signals[21]. However, whether a bloom produces toxins or is harmful to humans or the marine environment is not distinguishable from satellite data. We delineated bloom-affected areas (that is, the areas where algal blooms were detected), and enumerated the bloom count at the 1-km pixel level (that is, the number of detected blooms per pixel) (Fig. 1). We further estimated the bloom frequency (dimensionless) by integrating the bloom count and affected areas within 1° × 1° grid cells (see Methods), and this metric was used to examine temporal dynamics in bloom intensity. Validation with independent satellite samples selected via several visual inspection techniques showed an overall accuracy level of more than 95% for our method, and comparisons using discrete events in HAEDAT[6] indicated that we successfully identified bloom

counts for 79.3% of the historical HAB events in that database (Extended Data Figs. 2–6). We examined phytoplankton blooms in the exclusive economic zones (EEZs) of 153 coastal countries and in 54 large marine ecosystems (LMEs) (Extended Data Fig. 7). Our study area encompasses global continental shelves and outer margins of coastal currents, which offer the majority of marine resources available for human use (see Methods). Out of the 153 coastal countries examined, 126 were observed to have phytoplankton blooms (Fig. 1). The total bloom-affected area was 31.47 million km^2, equivalent to approximately 24.2% of the global land area and 8.6% of the global ocean area, with a median bloom count of 4.3 per year during the past 2 decades (Fig. 1b). Europe (9.52 million km^2—30.3% of the total affected area) and North America (6.78 million km^2—21.5% of the total affected area) contributed the largest bloom areas. By contrast, the most frequent blooms were found around Africa and South America (median bloom counts of more than 6.3 per year). Australia experienced the lowest frequency (2.4 per year) and affected area (2.84 million km^2—9.0% of the total affected area) of blooms.

Fig. 1: Global patterns of coastal phytoplankton blooms between 2003 and 2020.



a, The spatial distribution of annual mean bloom count based on daily satellite detections. **b**, Continental and global statistics for annual mean bloom count (South America (SA), *n* = 3,846,125; Africa (AF), *n* = 2,516,225; Europe (EU), *n* = 17,703,949; North America (NA), *n* = 10,034,286; Asia (AS), *n* = 5,371,158; Australia (AU), *n* = 2,781,998 pixel observations). The centre line represents the median value, bottom and top bounds of boxes are first and third quartiles, and the whiskers show a maximum of 1.5 times the interquartile range. **c**, Continental statistics for the long-term annual mean of bloom-affected areas (*n* = 18 years). The percentages show the corresponding contributions to the global total. The bars represent s.d. Open circles are the affected areas during different years. Map created using Python 3.8.

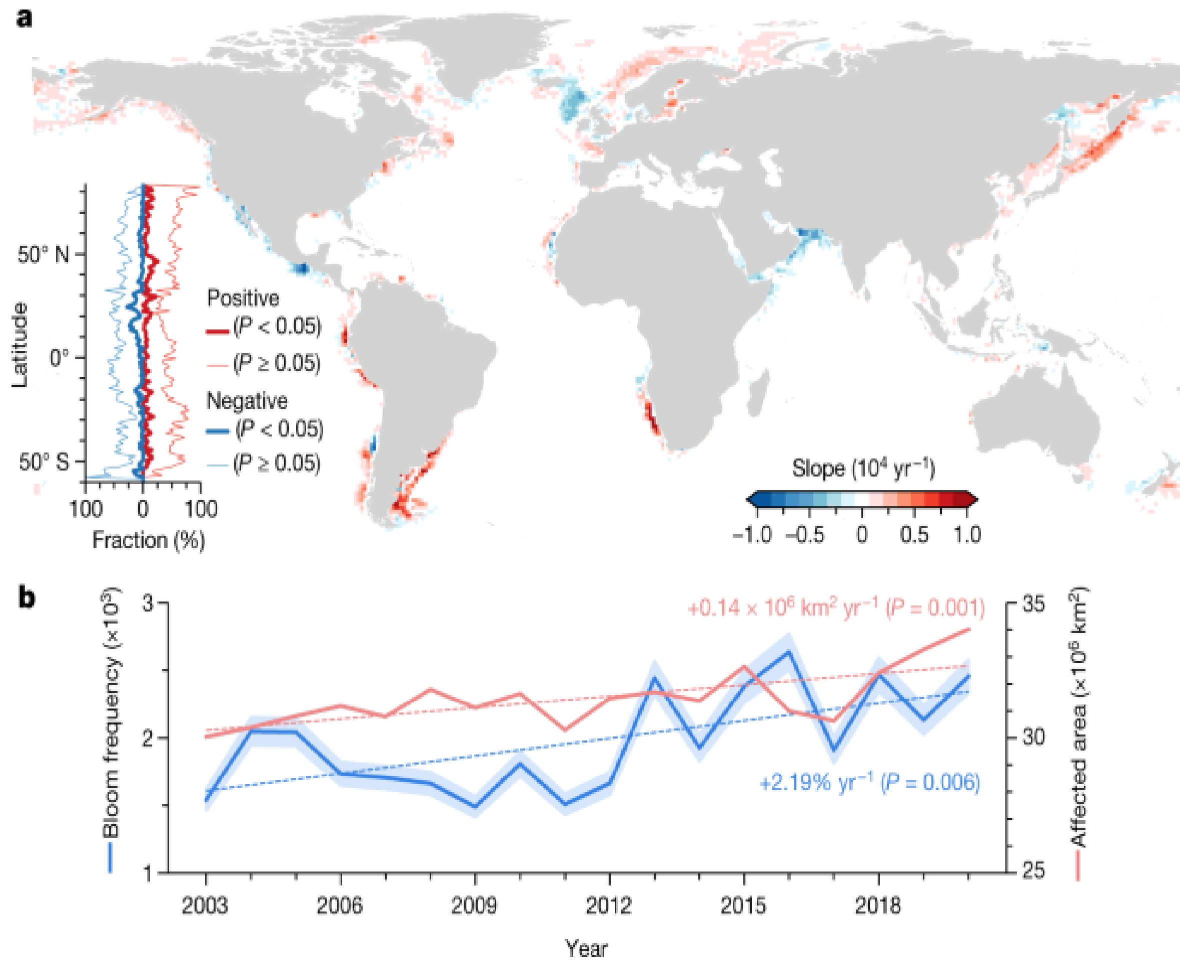
Source Data

Phytoplankton blooms occurred frequently in the eastern boundary current systems (that is, California, Benguela, Humboldt and Canary), northeastern USA, Latin America, the Baltic Sea, Northern Black Sea and the Arabian Sea (Fig. 1a). Five LMEs were found with the most frequent blooms (annual median bloom count over 15), including Patagonian Shelf, Northeast US continental Shelf, the Baltic Sea, Gulf of California and Benguela Current (Extended Data Fig. 7). These hotspots are often reported as having a high incidence of algal blooms, some of which are HABs, driven by either coastal upwelling or pronounced anthropogenic nutrient enrichment[9,22,23,24,25,26]. European LMEs showed mostly large proportions of bloom-affected areas, and some also showed frequent bloom occurrences. By contrast, Asian LMEs exhibited mainly infrequent blooms, given their large affected areas. We identified more bloom events in estuarine regions than along coasts in regions without major river discharge ($P < 0.05$; Extended Data Fig. 8), highlighting the critical role of terrestrial nutrient sources in coastal algal blooms[3].

Long-term trends

The total global bloom-affected area has expanded by 3.97 million km^2 (13.2%) between 2003 and 2020, equivalent to 0.14 million km^2 yr^{-1} ($P <$ 0.05; Fig. 2). Furthermore, the number of countries with significant bloom expansion was about 1.6 times those with a decreasing trend. The global median bloom frequency showed an increasing rate of 59.2% (+2.19% yr^{-1}, $P < 0.05$) over the observed period. Spatially, areas showing significant increasing trends ($P < 0.05$) in bloom frequency were 77.6% larger than those with the opposite trends (Fig. 2). Globally, our analysis revealed overall consistent fluctuations between the bloom-affected area and bloom frequency between 2003 and 2020 (Fig. 2b). However, there was no significant relationship between bloom extent and frequency in 23 countries and 10 LMEs over the past two decades, underscoring the spatial and temporal variability of algal blooms and the importance of continuous satellite monitoring.

Fig. 2: Trends of global coastal phytoplankton blooms between 2003 and 2020.



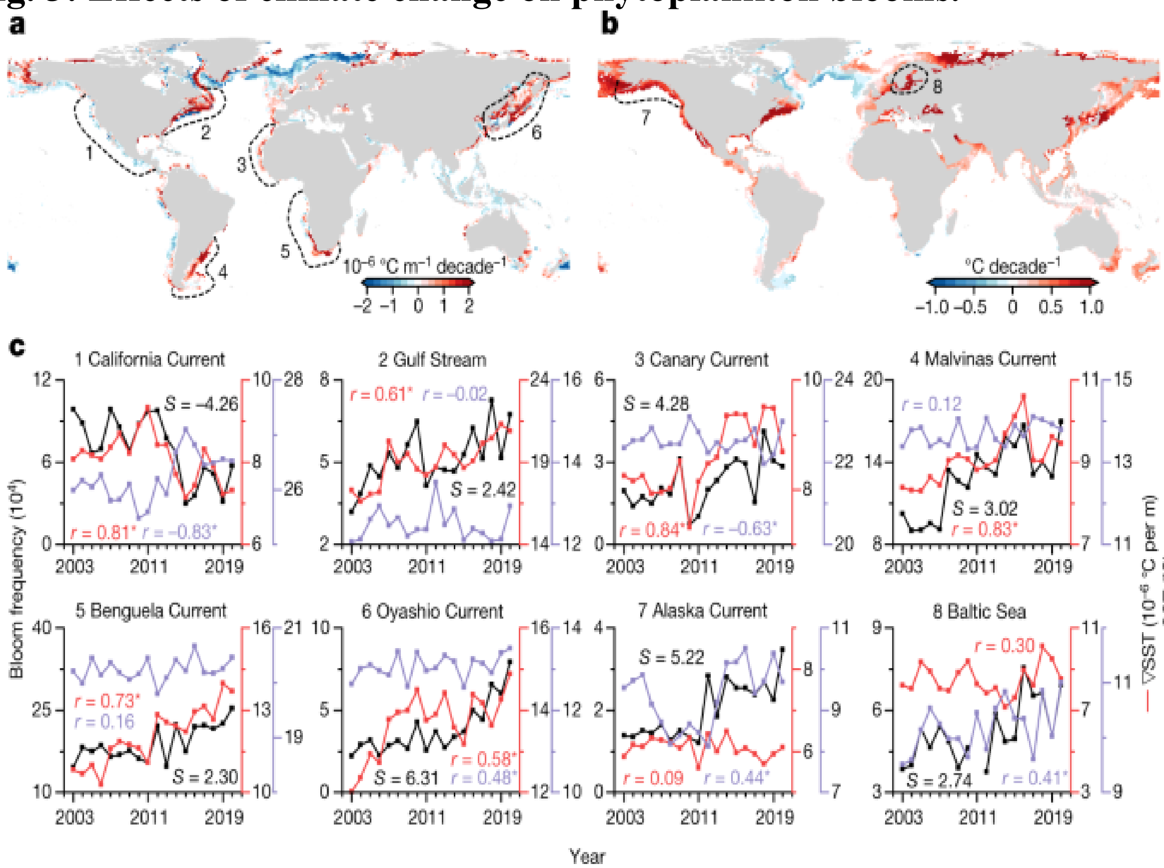
a, Spatial patterns of the trends in bloom frequency at a 1° × 1° grid scale. The latitudinal profiles show the fractions of grids with significant and insignificant trends (positive or negative) along the east–west direction. **b**, Interannual variability and trends in annual median bloom frequency and total global bloom-affected area. The linear slopes and *P*-value (two-sided *t*-test) are indicated. The shading associated with the bloom frequency data represents an uncertainty level of 5% in bloom detection. Map created using Python 3.8.

Source Data

The entire Southern Hemisphere was primarily characterized by increased bloom frequency, although weakened blooms were also sometimes found. In the Northern Hemisphere, the low latitude (<30° N) coasts were mainly featured with strong bloom weakening (Fig. 2a), primarily in the California Current System and the Arabian Sea. Bloom strengthening was found in the

northern Gulf of Mexico and the East and South China Seas, albeit at smaller magnitudes. At higher latitudes, weakening blooms were detected mainly in the northeastern North Atlantic and the Okhotsk Sea in the northwestern North Pacific. Globally, the largest increases in bloom frequency were observed in six major coastal current systems, including Oyashio (+6.31% yr^{-1}), Alaska (+5.22% yr^{-1}), Canary (+4.28% yr^{-1}), Malvinas (+3.02% yr^{-1}), Gulf Stream (+2.42% yr^{-1}) and Benguela (+2.30% yr^{-1}) (Figs. 2a and 3).

Fig. 3: Effects of climate change on phytoplankton blooms.



a,**b**, Global patterns of trends in SST gradient (**a**) and SST (**b**) from 2003 to 2020. **c**, Long-term changes in bloom frequency in the regions labelled in **a** and **b**, and their relationship to the SST and SST gradient. Linear slope (*S*) of bloom frequency and the correlation coefficient (*r*) between bloom frequency and the SST and the SST gradient (∇SST) are shown. Asterisks indicate statistically significant ($P < 0.05$) correlations. Maps created using ArcMap 10.4 and Python 3.8.

Source Data

Natural and anthropogenic effects

Increases in sea surface temperature (SST) can stimulate bloom occurrence. We found significant positive correlations ($P < 0.05$) between the annual mean bloom frequency and the coincident SST (SST data were averaged over the growth window of algal blooms within a year (Methods and Extended Data Fig. 9)) in several high-latitude regions (>40° N), such as the Alaska Current ($r = 0.44$), the Oyashio Current ($r = 0.48$) and the Baltic Sea ($r = 0.41$) (Fig. 3). These findings agree with previous studies, in which the bloom-favourable seasons in these temperate seas have been extended under warmer temperatures[27,28,29]. However, this temperature-based mechanism did not apply to regions with inconsistent trends between SST and bloom frequency, particularly for the substantial bloom weakening in the tropical and subtropical areas (Figs. 2a and 3b).

Changes in climate can also affect ocean circulation, altering ocean mixing and the transport of nutrients that drive the growth of marine phytoplankton and bloom formation[30,31,32]. We used the spatial SST gradient (in °C m^{-1}) as a proxy for the magnitude of oceanic mesoscale currents (the time-varying velocity of kinetic energy (also known as the eddy kinetic energy (EKE))) by following the methods of a previous study[33], and examined its effects on algal blooms (Methods). The trend in the SST gradient appeared more spatially aligned to bloom frequency than SST. We found significant positive correlations ($P < 0.05$) between the SST gradient and bloom frequency for various coastal current systems, including the Canary ($r = 0.84$), Malvinas ($r = 0.83$), California ($r = 0.81$), Benguela ($r = 0.73$), Gulf Stream ($r = 0.61$) and Oyashio ($r = 0.58$) currents.

Trends in bloom frequency in subtropical eastern boundary upwelling regions (the California, Benguela and Canary currents) followed the changes in mesoscale currents (Fig. 3a,c). In the California Current System, the decrease in bloom frequency was probably due to the weakened upwelling (represented by a reduced SST gradient and increased SST) and thus lower nutrient supply[25]. Conversely, the Canary and Benguela currents were characterized by strengthened upwelling and increased bloom frequency.

The two western boundary current systems at high latitudes (Malvinas and Oyashio)—although characterized by less pronounced upwelling[34]—exhibited a similar mechanism to the subtropical eastern boundary regions. For the subtropical western boundary Gulf Stream current, the strengthened current jets (manifested as a larger SST gradient) brought more nutrients from the continental shelf[35], triggering more algal blooms. Nevertheless, whether these changes in oceanic mesoscale activities were responses to wind, stratification, the shear of ocean currents or other factors[33] requires region-based investigations.

Global climate events, represented as the multivariate El Niño–Southern Oscillation index[36] (MEI), also showed connections with coastal bloom frequency. The minimum MEI in 2010 (a strong La Niña year) was followed by a low bloom frequency in the following year, and the largest MEI in 2015 (a strong El Niño year) was followed by the strongest bloom frequency in 2016 (Fig. 2b and Extended Data Fig. 10a).

Changes in anthropogenic nutrient enrichment may have also contributed to the trends in blooms[37]. For example, the decline in bloom frequency in the Arabian Sea, without clear links to SST or SST gradient changes, could result from decreased fertilizer use in the surrounding countries (such as Iran) (Extended Data Fig. 10). By contrast, the bloom strengthening in some Asian countries could be attributed to surges in fertilizer use[38]. We examined trends in fertilizer usage (either nitrogen or phosphorus) and bloom frequency and found high positive correlations in China, Iran, Vietnam and the Philippines. Paradoxically, decreased fertilizer uses and increased bloom frequency were identified in some countries, suggesting that nutrient control efforts might have been counterbalanced by the stimulatory effects of climate warming or other factors. Furthermore, the intensified aquaculture industry in Finland, China, Algeria, Guinea, Vietnam, Argentina, Russia and Uruguay may also be linked to their increased bloom incidence, as suggested by the significant positive correlations ($r > 0.5$, $P < 0.05$) between their aquaculture production and bloom frequency. A similar relationship between aquaculture expansion and positive trends in HAB incidence was reported from an analysis of HAEDAT data[6]. However, analogous positive feedbacks for fertilizer or aquaculture were not found in many other countries. Thus, an ecosystem model incorporating terrestrial and oceanic nutrient transport and

nutrient–plankton relationships of different species[39] is required to quantify the contributions of natural and anthropogenic factors to algal blooms[14].

Future implications

We acknowledge that our criteria for a detectable bloom event is operationally defined by sensor sensitivities and other factors, and that the bloom count metric used here may underestimate algal bloom incidence, particularly compared to harmful events entered in HAEDAT. For example, in a recent global analysis of the HAEDAT events, Hallegraeff et al.[6] report a dozen or more events per year for each of nine regions over a 33-year study period, compared to the global median bloom count of 4.3 in this study. There are several possible explanations for this discrepancy, such as the many low-cell-concentration HABs that are not detectable from space but that can still cause harm, as well as sensor sensitivities and algorithm thresholds. Furthermore, our bloom count was averaged over all 1-km pixels within the EEZs, whereas HAEDAT entries are based on discrete sampling regions. This underestimation does not, however, alter the trends and other conclusions of this study, as the metrics used here were constant across time and space. Underestimates would have been uniform across regions globally. In this regard, it is of note that the study of Hallegraeff et al.[6] found a significant link between the number of HAEDAT events over time and the global expansion of aquaculture production, similar to findings in our study.

The major contribution of our study is to provide a spatially and temporally consistent characterization of global coastal algal blooms between 2003 and 2020. Globally, increasing trends in algal bloom area and frequency are apparent. Regionally, however, trends were non-uniform owing to the compounded effects of changes in climate (such as changes in SST and SST gradient and climate extremes), anthropogenic eutrophication and aquaculture development. Our daily mapping of bloom events offers valuable baseline information to understand the mechanisms underlying the formation, maintenance, and dissipation of algal blooms[5,40]. This could aid in developing forecasting models (on either global or regional scales) that can help minimize the consequences of harmful blooms, and can also help in policy decisions relating to the control of nutrient discharges and other HAB-stimulatory factors. Noting again that many blooms are beneficial,

particularly in terms of their positive effects on ecosystems as well as on wild and farmed fisheries, the results here can also contribute toward policies and management actions that sustain those beneficial blooms.

Methods

Data sources

MODIS on the Aqua satellite provides a global coverage within 1–2 days. All images acquired by this satellite mission from January 2003 to December 2020 were used in our study to detect global coastal phytoplankton blooms, with a total of 0.76 million images. MODIS Level-1A images were downloaded from the Ocean Biology Distributed Active Archive Center (OB.DAAC) at NASA Goddard Space Flight Center (GSFC), and were subsequently processed with SeaDAS software (version 7.5) to obtain Rayleigh-corrected reflectance (R_{rc} (dimensionless), which was converted using the rhos (in sr^{-1}) product (rhos × π) from SeaDAS)[41], remote sensing reflectance (R_{rs} (sr^{-1})) and quality control flags (l2_flags). If a pixel was flagged by any of the following, it was then removed from phytoplankton bloom detection: straylight, cloud, land, high sunglint, high solar zenith angle and high sensor zenith angle (https://oceancolor.gsfc.nasa.gov/atbd/ocl2flags/). MODIS level-3 product for aerosol optical thicknesses (AOT) at 869 nm was also obtained from OB.DAAC NASA GSFC (version R2018.0), which was used to examine the impacts of aerosols on bloom trends.

We examined the algal blooms in the EEZs of 153 ocean-bordering countries (excluding the EEZs in the Caspian Sea or around the Antarctic), 126 of which were found with at least one bloom in the past two decades. The EEZ dataset is available at https://www.marineregions.org/download_file.php?name=World_EEZ_v11_20191118.zip. The EEZs are up to 200 nautical miles (or 370 km) away from coastlines, which include all continental shelf areas and offer the majority of marine resources available for human use. Regional statistics of algal blooms were also performed for LMEs. LMEs encompass global coastal oceans and outer edges of coastal currents areas, which are defined by various distinct features of the oceans, including

hydrology, productivity, bathymetry and trophically dependent populations[42]. Of the 66 LMEs identified globally, we excluded the Arctic and Antarctic regions and examined 54 LMEs. The boundaries of LMEs were obtained from https://www.sciencebase.gov/catalog/item/55c77722e4b08400b1fd8244.

We used HAEDAT to validate our satellite-detected phytoplankton blooms in terms of presence or absence. The HAEDAT dataset (http://haedat.iode.org) is a collection of records of HAB events, maintained under the UNESCO Intergovernmental Oceanographic Commission and with data archives since 1985. For each HAB event, the HAEDAT records its bloom period (ranging from days to months) and geolocation. We merged duplicate entries when both the recorded locations and times of the HAEDAT events were very similar to one another, and a total number of 2,609 HAEDAT events were ultimately selected between 2003 and 2020.

We used the ¼° resolution National Oceanic and Atmospheric Administration Optimum Interpolated SST (v. 2.1) data to examine the potential simulating effects of warming on the global phytoplankton trends. We also estimated the SST gradients following the method of Martínez-Moreno[33]. As detailed in ref. [33], the SST gradient can be used as a proxy for the magnitude of oceanic mesoscale currents (EKE). We used the SST gradient to explore the effects of ocean circulation dynamics on algal blooms.

Fertilizer uses and aquaculture production for different countries was used to examine the potential effects of nutrient enrichment from humans on global phytoplankton bloom trends. Annual data between 2003 and 2019 on synthetic fertilizer use, including nitrogen and phosphorus, are available from https://ourworldindata.org/fertilizers. Annual aquaculture production includes cultivated fish and crustaceans in marine and inland waters, and sea tanks, and the data between 2003 and 2018 are available from https://ourworldindata.org/grapher/aquaculture-farmed-fish-production.

The MEI, which combines various oceanic and atmospheric variables[36], was used to examine the connections between El Niño–Southern Oscillation activities and marine phytoplankton blooms. The dataset is available from https://psl.noaa.gov/enso/mei/.

Development of an automated bloom detection method

A recent study by the UNESCO Intergovernmental Oceanographic Commission revealed that globally reported HAB events have increased[6]. However, such an overall increasing trend was found to be highly correlated with recently intensified sampling efforts[6]. Once this potential bias was accounted for by examining the ratio between HAB events to the number of samplings[5], there was no significant global trend in HAB incidence, though there were increases in certain regions. With synoptic, frequent, and large-scale observations, satellite remote sensing has been extensively used to monitor algal blooms in oceanic environments[17,18,19]. For example, chlorophyll *a* (Chl*a*) concentrations, a proxy for phytoplankton biomass, has been provided as a standard product by NASA since the proof-of-concept Coastal Zone Color Scanner (1978–1986) era[43,44]. The current default algorithm used to retrieve Chl*a* products is based on the high absorption of Chl*a* at the blue band[45,46], which often shows high accuracy in the clear open oceans but high uncertainties in coastal waters. This is because, in productive and dynamic coastal oceans, the absorption of Chl*a* in the blue band can be obscured by the presence of suspended sediments and/or coloured dissolved organic matter (CDOM)[47]. To address this problem, various regionalized Chl*a* algorithms have been developed[48]. Unfortunately, the concentrations of the water constituents (CDOM, sediment and Chl*a*) can vary substantially across different coastal oceans. As a result, a universal Chl*a* algorithm that can accurately estimate Chl*a* concentrations in global coastal oceans is not currently available.

Alternatively, many spectral indices have been developed to identify phytoplankton blooms instead of quantifying their bloom biomass, including the normalized fluorescence line height[21] (nFLH), red tide index[49] (RI), algal bloom index[47] (ABI), red–blue difference (RBD)[50], *Karenia brevis* bloom index[50] (KBBI) and red tide detection index[51] (RDI). In practice, the most important task for these index-based algorithms is to determine their optimal thresholds for bloom classification. However, such optimal thresholds can be regional-or image-specific[20], due to the complexity of optical features in coastal waters and/or the contamination of unfavourable

observational conditions (such as thick aerosols, thin clouds, and so on), making it difficult to apply spectral-index-based algorithms at a global scale.

To circumvent the difficulty in determining unified thresholds for various spectral indices across global coastal oceans, an approach from a recent study to classify algal blooms in freshwater lakes[52] was adopted and modified here. In that study, the remotely sensed reflectance data in three visible bands (red, green and blue) were converted into two-dimensional colour space created by the Commission Internationale del'éclairage (CIE), in which the position on the CIE chromaticity diagram represented the colour perceived by human eyes (Extended Data Fig. 1a). As the algal blooms in freshwater lakes were manifested as greenish colours, the reflectance of bloom-containing pixels was expected to be distributed in the green gamut of the CIE chromaticity diagram; the stronger the bloom, the closer the distance to the upper border of the diagram (the greener the water).

Here, the colour of phytoplankton blooms in the coastal oceans can be greenish, yellowish, brownish, or even reddish[53], owing to the compositions of bloom species (diatoms or dinoflagellates) and the concentrations of different water constituents. Furthermore, the Chl*a* concentrations of the coastal blooms are typically lower than those in inland waters, thus demanding more accurate classification algorithms. Thus, the algorithm proposed by Hou et al.[52] was modified when using the CIE chromaticity space for bloom detection in marine environments. Specifically, we used the following coordinate conversion formulas to obtain the *xy* coordinate values in the CIE colour space:

$$\begin{array}{c}x=X/(X+Y+Z)\\ y=Y/(X+Y+Z)\\ X=2.7689R+1.7517G+1.1302B\\ Y=1.0000R+4.5907G+0.0601B\\ Z=0.0000R+0.0565G+5.5943B\end{array}$$
(1)

where R, G and B represent the R_{rc} at 748 nm, 678 nm (fluorescence band) and 667 nm in the MODIS Aqua data, respectively. By contrast, the R, G and B channels used in Hou et al.[52] were the red, green and blue bands. We used the fluorescence band for the G channel because, for a given region, the 678

nm signal increases monotonically with the Chl*a* concentration for blooms of moderate intensity[21], which is similar to the response of greenness to freshwater algal blooms. Thus, the converted *y* value in the CIE coordinate system represents the strength of the fluorescence. In practice, for pixels with phytoplankton blooms, the converted colours in the chromaticity diagram will be located within the green, yellow or orange–red gamut (see Extended Data Fig. 1a); the stronger the fluorescence signal is, the closer the distance to the upper border of the CIE diagram (larger *y* value). By contrast, for bloom-free pixels without a fluorescence signal, their converted *xy* coordinates will be located in the blue or purple gamut. Therefore, we can determine a lower boundary in the CIE two-dimensional coordinate system to separate bloom and non-bloom pixels, similar to the method proposed by Hou et al.[52].

We selected 53,820 bloom-containing pixels from the MODIS R_{rc} data as training samples to determine the boundary of the CIE colour space. These sample points were selected from nearshore waters worldwide where frequent phytoplankton blooms have been reported (Extended Data Fig. 2); the algal species included various species of dinoflagellates and diatoms[20]. A total of 80 images was used, which were acquired from different seasons and across various bloom magnitudes, to ensure that the samples used could almost exhaustively represent the different bloom conditions in the coastal oceans.

We combined the MODIS FLH_{Rrc} (fluorescence line height based on R_{rc}) and enhanced red–green–blue composite (ERGB) to delineate bloom pixels manually. The FLH_{Rrc} image was calculated as:

$$\begin{array}{c}{{\rm{FLH}}}_{{\rm{Rrc}}}={R}_{{\rm{rc}}678}\times {F}_{678}-[{R}_{{\rm{rc}}667}\times {F}_{667}+({R}_{{\rm{rc}}748}\times {F}_{748}\\ \,\,-\,{R}_{{\rm{rc}}667}\times {F}_{667})\times (678-667)/(748-667)]\end{array}$$
(2)

where R_{rc667}, R_{rc678} and R_{rc748} are the R_{rc} at 667, 678 and 748 nm, respectively, and F_{667}, F_{678} and F_{748} are the corresponding extraterrestrial

solar irradiance. ERGB composite images were generated using R_{rc} of three bands at 555 (R), 488 (G) and 443 nm (B). Although phytoplankton-rich and sediment-rich waters have high FLH_{Rrc} values, they appear as darkish and bright features in the ERGB images (Extended Data Fig. 3), respectively[21]. In fact, visual examination with fluorescence signals and ERGB has been widely accepted as a practical way to delineate coastal algal blooms on a limited number of images[21,54,55]. Note that the FLH_{Rrc} here was slightly different from the NASA standard nFLH product[56], as the latter is generated using R_{rs} (corrected for both Rayleigh and aerosol scattering) instead of R_{rc} (with residual effects of aerosols). However, when using the NASA standard algorithm to further perform aerosol scattering correction over R_{rc}, 20.7% of our selected bloom-containing pixels failed to obtain valid R_{rs} (without retrievals or flagged as low quality), especially for those with strong blooms (see examples in Extended Data Fig. 4). Likewise, we also found various nearshore regions with invalid R_{rs} retrievals. By contrast, R_{rc} had valid data for all selected samples and showed more coverage in nearshore coastal waters. The differences between R_{rs} and R_{rc} were because the assumptions for the standard atmospheric correction algorithm do not hold for bloom pixels or nearshore waters with complex optical properties[57]. In fact, R_{rc} has been used as an alternative to R_{rs} in various applications in complex waters[58,59].

We converted the R_{rc} data of 53,820 selected sample pixels into the *xy* coordinates in the CIE colour space (Extended Data Fig. 1a). As expected, these samples of bloom-containing pixels were located in the upper half of the chromaticity diagram (the green, yellow and orange–red gamut) (Extended Data Fig. 1a). We determined the lower boundary of these sample points in the chromaticity diagram, which represents the lightest colour and thus the weakest phytoplankton blooms; any point that falls above this boundary represents stronger blooms. The method to determine the boundary was similar to Hou et al.[52]: we first binned the sample points according to the *x* value in the chromaticity diagram and estimated the 1st percentile (Q_1%) of the corresponding *Y* for each bin; then, we fit the Q_1% using two-order polynomial regression. Sensitivity analysis with $Q_{0.3}$% (the three-sigma value) resulted in minor changes (<1%) in the resulting bloom

areas for single images. Notably, sample points were rarely located near white points ($x = 1/3$ and $y = 1/3$, represent equal reflection from three RGB bands) in the diagram, and we used two polynomial regressions to determine the boundaries for x values greater and less than 1/3, which can be expressed as:

$${y}_{1}=4.8093{x}^{2}-3.0958x+0.8357\,x < \frac{1}{3}$$
(3)

$${y}_{2}=4.9040{x}^{2}-3.5759x+0.9862\,x > \frac{1}{3}$$
(4)

Based on the above, if a pixel's *xy* coordinate (converted from R_{rc} spectrum) satisfies the conditions of ($x < 1/3$ AND $y > y_1$) or ($x > 1/3$ AND $y > y_2$), it is classified as a 'bloom' pixel.

Depending on the local region and application purpose, the meaning of 'phytoplankton bloom' may differ. Here, for a global application, the pixelwise bloom classification is based on the relationship (represented using the CIE colour space) between R_{rc} in the 667-, 678- and 754-nm bands derived from visual interpretation of the 80 pairs of FLH_{Rrc} and ERGB imagery. Instead of a simple threshold, we used a lower boundary of the sample points in the chromaticity diagram to define a bloom. In simple words, a pixel is classified as a bloom if its fluorescence signal is detectable (the associated *xy* coordinate in the CIE colour space located above the lower boundary). Histogram of the nFLH values from the 53,820 training pixels demonstrated the minimum value of ~0.02 mW cm^{-2} μm^{-1} (Extended Data Fig. 1a), which is in line with the lower-bound signal of *K. brevis* blooms on the West Florida shelf[21,47]. Note that, such a minimum nFLH is determined from the global training pixels, and it does not necessarily represent a unified lower bound for phytoplankton blooms across the entire globe, especially considering that fluorescence efficiency may be a large variable across different regions. Different regions may have different lower bounds of nFLH to define a bloom, and such variability is represented by the predefined boundary in the CIE chromaticity diagram in our study. Correspondingly, although the accuracy of Chl*a* retrievals may have large uncertainties in coastal waters, the histogram of the 53,820 training pixels shows a lower bound of ~1 mg m^{-3} (Extended Data Fig. 1a). Similarly to

nFLH, such a lower bound may not be applicable to all coastal regions, as different regions may have different lower bounds of Chl*a* for bloom definition.

Although the MODIS cloud (generated by SeaDAS with $R_{rc869} < 0.027$) and associated straylight flags can be used to exclude most clouds, we found that residual errors from thin clouds or cloud shadows could affect the spectral shape and cause misclassification for bloom detections. Thus, we designed two spectral indices to remove such effects:

$${\rm{Index}}1={\rm{n}}{R}_{{\rm{rc}}488}-{{\rm{n}}R}_{{\rm{rc}}443}-({{\rm{n}}R}_{{\rm{rc}}555}-{{\rm{n}}R}_{{\rm{rc}}443})\times 0.5$$
(5)
$${\rm{Index}}2={{\rm{n}}R}_{{\rm{rc}}555}-{{\rm{n}}R}_{{\rm{rc}}469}-({{\rm{n}}R}_{{\rm{rc}}645}-{{\rm{n}}R}_{{\rm{rc}}469})\times 0.5$$
(6)

where Index1 and Index2 were used to remove cloud shadows and clouds, respectively. The nR_{rc443}, nR_{rc488} and nR_{rc555} in index1 are the normalized R_{rc}, obtained by normalizing R_{rc488}. Similar calculations were performed for index2. The purpose of normalizations is to eliminate the effect of the absolute magnitude of the reflectance, so that the thresholds of these two indices are influenced by only the relative magnitude (spectral shape). We determined thresholds for Index1 (>0.12) and Index2 (<0.012) through trial-and-error and ensured that the misclassifications caused by residual errors from clouds and cloud shadows could be effectively removed. After applying the cloud/cloud shadow and various other masks that are associated with l2_flags, we obtained an annual mean valid pixel observation (N_{vobs}) of ~2.0×10^5 for global 1° × 1° grid cells, and the fluctuation patterns and trends of N_{vobs}, either annually or seasonally, are different from that of the global bloom frequency and affected areas (see Supplementary Fig. 1).

Assessments of the algorithm performance

In addition to phytoplankton blooms, macroalgal blooms (*Sargassum* and *Ulva*) frequently occur in many coastal oceans[60,61,62,63]. To verify whether our CIE-fluorescence algorithm could eliminate such impacts, we compared the spectra between micro-and macroalgal blooms (see Extended Data Fig. 1b). We found that the spectral shapes of *Sargassum* and *Ulva* are substantially different from those of microalgae, particularly for the three bands used for CIE coordinate conversion. The converted *xy* coordinates for macroalgae were located in the purple–red gamut of the CIE diagram, which was far below the predefined boundary (Extended Data Fig. 1). Moreover, our algorithm is not affected by highly turbid waters for the following two reasons: first, extremely high turbidity tends to saturate the MODIS ocean bands[64], which can be easily excluded; second, without a fluorescence peak, the reflectance of unsaturated turbid waters, after conversion to CIE coordinates, will be located below the predefined boundary (see example in Extended Data Fig. 1b). We also confirmed that the spectral shapes of coccolithophore blooms are different from dinoflagellates and diatoms (see example in Extended Data Fig. 1b), and thus they are excluded from our algorithm.

Three different types of validation methods were adopted to demonstrate the reliability of the proposed CIE-fluorescence algorithm for phytoplankton bloom detection in global coastal oceans, including visual inspections of the RGB, ERGB and FLH_{Rrc} images, verifications using independent manually delineated algal blooms, and comparisons with the reported HAB events from the HAEDAT dataset.

First, we selected MODIS Aqua images from different locations where coastal phytoplankton blooms have been recorded in the published literature. We visually compared the RGB, ERGB, and FLH_{Rrc} images, and our algorithm detected bloom patterns (see examples in Extended Data Fig. 3). Comparisons from various images worldwide showed that our algorithm could successfully identify regions with high FLH_{Rrc} values and brownish-to-darkish features on the ERGB images, which can be considered phytoplankton blooms.

Second, we delineated additional 15,466 bloom-containing pixels from 35 images covering different coastal areas, using the same visual inspection and

manual delineation method as for the training sample pixels. Moreover, we also selected 14,149 bloom-free pixels (bright or turquoise green colours on ERGB images or low FLH_{Rrc} values) within the same images as bloom-containing images. We applied our algorithm to all these pixels, and compared the algorithm-identified and manually delineated results. Our CIE-fluorescence algorithm showed high values in both producer and user accuracies (92.04% and 98.63%) (Supplementary Table 1), and appeared effective at identifying bloom pixels and excluding false negatives (blooms classified as non-blooms) and false positives (non-blooms classified as blooms).

Third, we validated the satellite-detected phytoplankton blooms using in situ reported HAB events from the HAEDAT dataset. For each HAB event in the HAEDAT dataset, we obtained all MODIS images over the reported bloom period (from days to months). Within each year, we estimated the ratio between the number of satellite images with 'bloom detected' against the number of valid images (see definition above) during the bloom periods across the entire globe (Supplementary Table 1). Moreover, we calculated the number of events with at least one successful satellite bloom detection (N_s), and then estimated the ratio between N_s and the total HAB events for each year. Results showed that substantial amounts (averaged at 51.2%) of satellite observations acquired during the HAB event periods were found with phytoplankton blooms by our algorithm. Overall, 79.3% of the global HAB events were successfully identified by satellite. The discrepancies between satellite and in situ observations could be explained by the following reasons: first, our study focused only on the phytoplankton blooms that are resolvable by satellite fluorescence signals; other types of HAB events in the HAEDAT dataset may not have been detectable by satellite observations, such as events with lower phytoplankton biomass but high toxicity, occurrences at the subsurface layers, or fluorescence signals overwhelmed by suspended sediments[65,66,67]. Second, although the HAEDAT recorded HAB events could be sustained for long periods, high biomass of surface algae may not have occurred every day within this period due to the changes in stratification, precipitation, wind, vertical migration of cells, and many other factors[68]. Third, the spatial scale of certain HAB events may have been too small to be identified using the 1-km resolution MODIS observations. Fourth, a reduced MODIS satellite observation

frequency by the contaminations of clouds and land adjacency effects[69]. Therefore, we believe the underestimations of satellite-detected blooms compared to the in situ reported HAB events were mainly due to inconsistencies between the two observations rather than uncertainties in our algorithm.

Because R_{rc} depends not only on water colour but also on aerosols (type and concentration) and solar and viewing geometry, a sensitivity analysis was used to determine whether such variables could impact bloom detection. Aerosol reflectance (ρ_a) with different AOTs at 869 nm was simulated using the NASA-recommended maritime aerosol model (r75f02, with a relative humidity of 75% and a fine-mode fraction of 2%). Then, ρ_a of each MODIS band was added to R_{rc} images, and the resulting bloom areas with and without added ρ_a were compared. Results showed that even with a change of 0.02 in AOT at 869 nm, the bloom areas showed minor changes (<2%) in the tested images; minor changes were also found when we used different aerosol models to conduct ρ_a simulations[70]. Note that 0.02 represents the high end of the AOT intra-annual variability in coastal oceans (see Extended Data Fig. 5), and the associated interannual changes are much smaller. Thus, the use of R_{rc} instead of the fully atmospherically corrected reflectance R_{rs} could have limited impacts on our detected global bloom trend.

We also tried various index-based algorithms developed in previous studies. However, results showed that all these methods require image-specific thresholds to accurately determine algal bloom boundaries for different coastal regions (see Extended Data Fig. 6). By contrast, although our CIE-fluorescence algorithm may lead to different bloom thresholds for different regions, it can identify bloom pixels without adjusting the coefficients and, therefore, is more suitable for global-scale bloom assessment efforts.

We acknowledge that our satellite-detected algal blooms represent only high amounts of phytoplankton biomass on the ocean surfaces without distinguishing whether such blooms produce toxins or are harmful to marine environments. Furthermore, with only limited spectral information from MODIS, it is difficult to discriminate the phytoplankton species of algal blooms; such information could help to improve our understanding of the

impacts of these phytoplankton blooms. However, we expect these challenges to be addressed soon with the scheduled launch of the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission by NASA in 2024, where the hyperspectral measurements over a broad spectrum (350–885 nm) will make species-level classifications possible[71].

Exploring the patterns and trends of global coastal phytoplankton blooms

We applied the CIE-fluorescence algorithm to all MODIS Aqua level-2 R_{rc} images, and a total number of 0.76 million images between 2003 and 2020 were processed. We mapped all detected blooms into 1-km daily scale level-3 composites. The number of bloom counts within a year for each location can be easily enumerated, and the long-term annual mean values were then estimated (Fig. 1a). We further calculated the total global bloom-affected area (the areas where algal blooms were detected at least once) for each year and examined their changes over time (Fig. 2b).

We defined bloom frequency (dimensionless) to represent the density of phytoplankton blooms for a year by integrating the bloom count and bloom-affected areas within 1°×1° grid cells within that year, which is expressed as:

$${\rm{Bloom}}\,{\rm{frequency}}=\frac{n}{N}\mathop{\sum }\limits_{i=1}^{n}{M}_{i}$$
(7)

where M_i is the enumerated bloom count for each 1-km resolution pixel in a year within one 1° × 1° grid cell, and n represents the associated number of bloom-affected pixels in the same cell (the number of pixels with $M_i > 0$), and N is the total number of 1-km MODIS pixels in this grid cell. We estimated the bloom frequency for each year between 2003 and 2020, and determined the long-term trend over global EEZs through a linear least-squares regression (see Fig. 2a).

Continental and country-level statistics were performed for bloom count, bloom-affected areas, and bloom frequency (Fig. 1b,c and Supplementary Table 2), using boundaries for the EEZs of different ocean-bordering

countries (see above). Similar statistics were also conducted for 54 LMEs (Extended Data Fig. 7 and Supplementary Table 3).

Correlations with SST and SST gradient

To assess the impacts of climate change on long-term trends in coastal phytoplankton blooms, we correlated the annual mean bloom frequency and the associated SST and SST gradient in various coastal current systems for grid cells with significant changes in bloom frequency (Fig. 3c). The SST and SST gradient were averaged over the growth window within a year, assuming that the changes within the growth window, either in water temperatures or ocean circulations, play more important roles in the bloom trends compared to other seasons[32].

We determined the growth window of phytoplankton blooms for each 1° × 1° grid cell (Extended Data Fig. 9a) using the following method: first, we estimated the proportion of cumulative bloom-affected pixels within the grid cells for a year. Second, a generalized additive model[72] was used to determine the shape of the phenological curves (Extended Data Fig. 9b), where a log link function and a cubic cyclic regression spline smoother were applied[73,74]. Third, the timing of maximum bloom-affected areas (TMBAA) was then determined by identifying the inflection point on the bloom growth curve (Extended Data Fig. 9c). To facilitate comparisons across Northern and Southern Hemispheres, the year in the Southern Hemisphere was shifted forward by 183 days (Extended Data Fig. 9c). We characterized the similarity of the bloom growth curve between different grid cells and grouped them into three distinct clusters using a fuzzy c-means cluster analysis method[75,76]. We found uniform distributions of the clusters over large geographic areas. Cluster I is mainly distributed in mid-low latitudes (<45° N and <30° S), where the maximum bloom-affected areas were expected in the early period of the year. Cluster II was mostly found in higher latitudes, with bloom developments (quasi-) synchronized with increases in SST. Cluster III was detected along the coastlines, where the bloom-affected areas increase throughout the entire year. In practice, the growth window for clusters I and III was set as the entire year, and that for cluster II was set from day 150 to day 270 within the year. We further found that the TMBAA for cluster II showed small changes over the entire period

(Extended Data Fig. 9d), indicating relatively stable phenological cycles for those phytoplankton blooms[32,77].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The satellite-based dataset of global coastal algal bloom at 1-km resolution and the associated code are available at https://doi.org/10.5281/zenodo.7359262. Source data are provided with this paper.

Change history

- **06 March 2023**

  In the version of this article initially published, there was a typo in the last term of equation (4), now corrected to read "$x > 1/3$," where the greater than symbol was missing.

References

1. Breitburg, D. et al. Declining oxygen in the global ocean and coastal waters. *Science* **359**, eaam7240 (2018).

2. Anderson, D. M. Turning back the harmful red tide. *Nature* **388**, 513–514 (1997).

3. Beman, J. M., Arrigo, K. R. & Matson, P. A. Agricultural runoff fuels large phytoplankton blooms in vulnerable areas of the ocean. *Nature* **434**, 211–214 (2005).

4. Heisler, J. et al. Eutrophication and harmful algal blooms: a scientific consensus. *Harmful Algae* **8**, 3–13 (2008).

5. Anderson, D. M., Cembella, A. D. & Hallegraeff, G. M. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annu. Rev. Mar. Sci.* **4**, 143–176 (2012).

6. Hallegraeff, G. M. et al. Perceived global increase in algal blooms is attributable to intensified monitoring and emerging bloom impacts. *Commun. Earth Environ.* **2**, 117 (2021).

7. Smith, V. H. Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environ. Sci. Pollut. Res.* **10**, 126–139 (2003).

8. Fleming, L. E. et al. Review of Florida red tide and human health effects. *Harmful Algae* **10**, 224–233 (2011).

9. Richlen, M. L., Morton, S. L., Jamali, E. A., Rajan, A. & Anderson, D. M. The catastrophic 2008–2009 red tide in the Arabian Gulf region, with observations on the identification and phylogeny of the fish-killing dinoflagellate *Cochlodinium polykrikoides*. *Harmful Algae* **9**, 163–172 (2010).

10. Hallegraeff, G. & Bolch, C. Unprecedented toxic algal blooms impact on Tasmanian seafood industry. *Microbiol. Aust.* **37**, 143–144 (2016).

11. Diaz, R. J. & Rosenberg, R. Spreading dead zones and consequences for marine ecosystems. *Science* **321**, 926–929 (2008).

12. Barton, A. D., Irwin, A. J., Finkel, Z. V. & Stock, C. A. Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. *Proc. Natl Acad. Sci. USA* **113**, 2964–2969 (2016).

13. Gobler, C. J. Climate change and harmful algal blooms: insights and perspective. *Harmful Algae* **91**, 101731 (2020).

14. Zohdi, E. & Abbaspour, M. Harmful algal blooms (red tide): a review of causes, impacts and approaches to monitoring and prediction. *Int. J. Environ. Sci. Technol.* **16**, 1789–1806 (2019).

15. Wells, M. L. et al. Future HAB science: directions and challenges in a changing climate. *Harmful Algae* **91**, 101632 (2020).

16. Rabalais, N. N., Turner, R. E., Díaz, R. J. & Justić, D. Global change and eutrophication of coastal waters. *ICES J. Mar. Sci.* **66**, 1528–1537 (2009).

17. Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R. & Brando, V. E. A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Prog. Oceanogr.* **123**, 123–144 (2014).

18. Wolny, J. L. et al. Current and future remote sensing of harmful algal blooms in the Chesapeake Bay to support the shellfish industry. *Front. Mar. Sci.* **7**, 337 (2020).

19. Stumpf, R. P. et al. Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. *Harmful Algae* **2**, 147–160 (2003).

20. Bernard, S., Kudela, R. M., Robertson Lain, L. & Pitcher, G. *Observation of Harmful Algal Blooms with Ocean Colour Radiometry* http://dx.doi.org/10.25607/OBP-1042 (IOCCG, 2021).

21. Hu, C. et al. Red tide detection and tracing using MODIS fluorescence data: a regional example in SW Florida coastal waters. *Remote Sens. Environ.* **97**, 311–321 (2005).

22. Andersen, J. H. et al. Long-term temporal and spatial trends in eutrophication status of the Baltic Sea. *Biol. Rev.* **92**, 135–149 (2017).

23. Gómez, F. & Boicenco, L. An annotated checklist of dinoflagellates in the Black Sea. *Hydrobiologia* **517**, 43–59 (2004).

24. Townsend, D. W., Pettigrew, N. R. & Thomas, A. C. Offshore blooms of the red tide dinoflagellate, *Alexandrium* sp., in the Gulf of Maine. *Cont. Shelf Res.* **21**, 347–369 (2001).

25. Pitcher, G. C., Figueiras, F. G., Hickey, B. M. & Moita, M. T. The physical oceanography of upwelling systems and the development of harmful algal blooms. *Prog. Oceanogr.* **85**, 5–32 (2010).

26. López-Cortés, D. J. et al. The state of knowledge of harmful algal blooms of *Margalefidinium polykrikoides* (a.k.a. *Cochlodinium polykrikoides*) in Latin America. *Front. Mar. Sci.* **6**, 463 (2019).

27. Anderson, D. M. et al. Evidence for massive and recurrent toxic blooms of *Alexandrium catenella* in the Alaskan Arctic. *Proc. Natl Acad. Sci. USA* **118**, e2107387118 (2021).

28. Griffith, A. W., Doherty, O. M. & Gobler, C. J. Ocean warming along temperate western boundaries of the Northern Hemisphere promotes an expansion of *Cochlodinium polykrikoides* blooms. *Proc. R. Soc. B* **286**, 20190340 (2019).

29. Conley, D. J. Save the Baltic Sea. *Nature* **486**, 463–464 (2012).

30. Mahadevan, A., D'Asaro, E., Lee, C. & Perry, M. J. Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms. *Science* **337**, 54–58 (2012).

31. Chelton, D. B., Gaube, P., Schlax, M. G., Early, J. J. & Samelson, R. M. The influence of nonlinear mesoscale eddies on near-surface oceanic chlorophyll. *Science* **334**, 328–332 (2011).

32. Boyce, D. G., Petrie, B., Frank, K. T., Worm, B. & Leggett, W. C. Environmental structuring of marine plankton phenology. *Nat. Ecol. Evol.* **1**, 1484–1494 (2017).

33. Martínez-Moreno, J. et al. Global changes in oceanic mesoscale currents over the satellite altimetry record. *Nat. Clim. Change* **11**, 397–403 (2021).

34. Kämpf, J. & Chapman, P. in *Upwelling Systems of the World* 31–65 (Springer, 2016).

35. Lee, T. N., Yoder, J. A. & Atkinson, L. P. Gulf Stream frontal eddy influence on productivity of the southeast US continental shelf. *J. Geophys. Res.* **96**, 22191–22205 (1991).

36. Wolter, K. & Timlin, M. S. El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext). *Int. J. Climatol.* **31**, 1074–1087 (2011).

37. Glibert, P. M. & Burford, M. A. Globally changing nutrient loads and harmful algal blooms: recent advances, new paradigms, and continuing challenges. *Oceanography* **30**, 58–69 (2017).

38. Lu, C. & Tian, H. Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hot spots and nutrient imbalance. *Earth Syst. Sci. Data* **9**, 181–192 (2017).

39. Falkowski, P. G., Barber, R. T. & Smetacek, V. Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**, 200–206 (1998).

40. Wells, M. L. et al. Harmful algal blooms and climate change: Learning from the past and present to forecast the future. *Harmful Algae* **49**, 68–93 (2015).

41. Hu, C. A novel ocean color index to detect floating algae in the global oceans. *Remote Sens. Environ.* **113**, 2118–2129 (2009).

42. Sherman, K. Adaptive management institutions at the regional level: the case of large marine ecosystems. *Ocean Coast. Manag.* **90**, 38–49 (2014).

43. Gordon, H. R., Clark, D. K., Mueller, J. L. & Hovis, W. A. Phytoplankton pigments from the Nimbus-7 coastal zone color scanner: comparisons with surface measurements. *Science* **210**, 63–66 (1980).

44. Moore, J. K. & Abbott, M. R. Phytoplankton chlorophyll distributions and primary production in the Southern Ocean. *J. Geophys. Res.* **105**, 28709–28722 (2000).

45. Hu, C. et al. Improving satellite global chlorophyll *a* data products through algorithm refinement and data recovery. *J. Geophys. Res.* **124**, 1524–1543 (2019).

46. Hu, C., Lee, Z. & Franz, B. Chlorophyll *a* algorithms for oligotrophic oceans: a novel approach based on three-band reflectance difference. *J. Geophys. Res.* **117**, C01011 (2012).

47. Hu, C. & Feng, L. Modified MODIS fluorescence line height data product to improve image interpretation for red tide monitoring in the eastern Gulf of Mexico. *J. Appl. Remote Sens.* **11**, 012003 (2016).

48. Siswanto, E. et al. Empirical ocean-color algorithms to retrieve chlorophyll-*a*, total suspended matter, and colored dissolved organic matter absorption coefficient in the Yellow and East China Seas. *J. Oceanogr.* **67**, 627–650 (2011).

49. Ahn, Y.-H. & Shanmugam, P. Detecting the red tide algal blooms from satellite ocean color observations in optically complex Northeast-Asia coastal waters. *Remote Sens. Environ.* **103**, 419–437 (2006).

50. Amin, R. et al. Novel optical techniques for detecting and classifying toxic dinoflagellate *Karenia brevis* blooms using satellite imagery. *Opt. Express* **17**, 9126–9144 (2009).

51. Shen, F., Tang, R., Sun, X. & Liu, D. Simple methods for satellite identification of algal blooms and species using 10-year time series data from the East China Sea. *Remote Sens. Environ.* **235**, 111484 (2019).

52. Hou, X. et al. Global mapping reveals increase in lacustrine algal blooms over the past decade. *Nat. Geosci.* **15**, 130–134 (2022).

53. Dierssen, H. M., Kudela, R. M., Ryan, J. P. & Zimmerman, R. C. Red and black tides: quantitative analysis of water-leaving radiance and

perceived color for phytoplankton, colored dissolved organic matter, and suspended sediments. *Limnol. Oceanogr.* **51**, 2646–2659 (2006).

54. Zhao, J., Temimi, M. & Ghedira, H. Characterization of harmful algal blooms (HABs) in the Arabian Gulf and the Sea of Oman using MERIS fluorescence data. *ISPRS J. Photogramm. Remote Sens.* **101**, 125–136 (2015).

55. Qi, L. et al. Noctiluca blooms in the East China Sea bounded by ocean fronts. *Harmful Algae* **112**, 102172 (2022).

56. Behrenfeld, M. J. et al. Satellite-detected fluorescence reveals global physiology of ocean phytoplankton. *Biogeosciences* **6**, 779–794 (2009).

57. Gordon, H. R. Atmospheric correction of ocean color imagery in the Earth Observing System era. *J. Geophys. Res.* **102**, 17081–17106 (1997).

58. Feng, L., Hou, X., Li, J. & Zheng, Y. Exploring the potential of Rayleigh-corrected reflectance in coastal and inland water applications: a simple aerosol correction method and its merits. *ISPRS J. Photogramm. Remote Sens.* **146**, 52–64 (2018).

59. Feng, L., Hu, C., Han, X., Chen, X. & Qi, L. Long-term distribution patterns of chlorophyll-*a* concentration in China's largest freshwater lake: MERIS full-resolution observations with a practical approach. *Remote Sens.* **7**, 275–299 (2015).

60. Xiao, J. et al. An anomalous bi-macroalgal bloom caused by *Ulva* and *Sargassum* seaweeds during spring to summer of 2017 in the western Yellow Sea, China. *Harmful Algae* **93**, 101760 (2020).

61. Teichberg, M. et al. Eutrophication and macroalgal blooms in temperate and tropical coastal waters: nutrient enrichment experiments with *Ulva* spp. *Glob. Change Biol.* **16**, 2624–2637 (2010).

62. Viaroli, P. et al. Nutrient and iron limitation to Ulva blooms in a eutrophic coastal lagoon (Sacca di Goro, Italy). *Hydrobiologia* **550**, 57–71 (2005).

63. Dierssen, H. M., Chlus, A. & Russell, B. Hyperspectral discrimination of floating mats of seagrass wrack and the macroalgae *Sargassum* in coastal waters of Greater Florida Bay using airborne remote sensing. *Remote Sens. Environ.* **167**, 247–258 (2015).

64. Hu, C. et al. Dynamic range and sensitivity requirements of satellite ocean color sensors: learning from the past. *Appl. Opt.* **51**, 6045–6062 (2012).

65. Trainer, V. L. et al. Pelagic harmful algal blooms and climate change: lessons from nature's experiments with extremes. *Harmful Algae* **91**, 101591 (2020).

66. Mardones, J. I. et al. Disentangling the environmental processes responsible for the world's largest farmed fish-killing harmful algal bloom: Chile, 2016. *Sci. Total Environ.* **766**, 144383 (2021).

67. Gilerson, A. et al. Fluorescence component in the reflectance spectra from coastal waters. II. Performance of retrieval algorithms. *Opt. Express* **16**, 2446–2460 (2008).

68. Lee, J. H., Harrison, P. J., Kuang, C. & Yin, K. in *The Environment in Asia Pacific Harbours* (ed. Wolanski, E.) 187–206 (Springer, 2006).

69. Feng, L. & Hu, C. Cloud adjacency effects on top-of-atmosphere radiance and ocean color data products: a statistical assessment. *Remote Sens. Environ.* **174**, 301–313 (2016).

70. Ahmad, Z. et al. New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans. *Appl. Opt.* **49**, 5545–5560 (2010).

71. Werdell, P. J. et al. The Plankton, Aerosol, Cloud, Ocean Ecosystem mission: status, science, advances. *Bull. Am. Meteorol. Soc.* **100**, 1775–1794 (2019).

72. Hastie, T. J. *Generalized Additive Models* (Routledge, 2017).

73. Wood, S. N. *Generalized Additive Models: An Introduction with R* (CRC press, 2017).

74. Macgregor, C. J. et al. Climate-induced phenology shifts linked to range expansions in species with multiple reproductive cycles per year. *Nat. Commun.* **10**, 4455 (2019).

75. Bezdek, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms* (Springer, 2013).

76. Bi, S. et al. Optical classification of inland waters based on an improved fuzzy C-means method. *Opt. Express* **27**, 34838–34856 (2019).

77. Kheireddine, M., Mayot, N., Ouhssain, M. & Jones, B. H. Regionalization of the Red Sea based on phytoplankton phenology: a satellite analysis. *J. Geophys. Res.* **126**, e2021JC017486 (2021).

Acknowledgements

We thank NASA for providing global MODIS satellite images, and the Intergovernmental Oceanographic Commission (IOC) of UNESCO for providing the HAEDAT dataset. L.F. was supported by the National Natural Science Foundation of China (no. 41890852, 42271322 and 41971304). D.M.A. was supported by the Woods Hole Center for Oceans and Human Health (National Science Foundation grant OCE-1840381 and National Institutes of Health grants NIEHS-1P01-ES028938-01).

Author information

Author notes

1. These authors contributed equally: Yanhui Dai, Shangbo Yang

Authors and Affiliations

1. School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China

   Yanhui Dai, Shangbo Yang, Dan Zhao, Luke Gibson, Chunmiao Zheng & Lian Feng

2. College of Marine Science, University of South Florida, St. Petersburg, FL, USA

   Chuanmin Hu

3. Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province, Shenzhen, China

   Wang Xu

4. Woods Hole Oceanographic Institution, Woods Hole, MA, USA

   Donald M. Anderson

5. School of Marine Science and Policy, College of Earth, Ocean, and Environment, University of Delaware, Lewes, DE, USA

   Yun Li

6. Department of Geographical Sciences, University of Maryland, College Park, MD, USA

   Xiao-Peng Song

7. Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, Nova Scotia, Canada

   Daniel G. Boyce

8. EIT Institute for Advanced Study, Ningbo, China

   Chunmiao Zheng

Contributions

Y.D. and S.Y.: methodology, data processing and analyses, and writing. L.F.: conceptualization, methodology, funding acquisition, supervision and writing. D.Z.: data processing and analysis. C.H., W.X., D.M.A., Y.L., X.-P.S., D.G.B., L.G. and C.Z. participated in interpreting the results and refining the manuscript.

Corresponding author

Correspondence to Lian Feng.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Bryan Franz, Bingkun Luo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Development of the CIE-fluorescence algorithm to detect phytoplankton blooms using MODIS

satellite imagery.

(**a**). A1: The density plot of manually delineated bloom-containing pixels in the CIE coordinate system (*n* = 53,820), and their distribution in the CIE color space (box in A2). A3: Histograms of nFLH and Chla for the delineated pixels, obtained using NASA standard algorithms[47,57]. (**b**) MODIS true color composites and selected spectra for phytoplankton blooms, macroalgal blooms (*Ulva* and *Sargassum*), coccolithophore blooms, and sediment-rich turbid waters. The x-y numbers indicate their corresponding positions in the CIE coordinate system. The black rectangular boxes in the three lower panels highlight different spectral shapes between phytoplankton blooms and other features near the fluorescence band. Maps created using ArcMap 10.4.

Source Data

Extended Data Fig. 2 MODIS-detected bloom count within certain years for several coastal regions with frequently reported blooms.

The MODIS observational year is annotated within each panel, and overlaid points indicate *in situ* recorded harmful algal bloom events from the Harmful Algae Event Database (HAEDAT) within the same year. The lower right panel shows the locations of all the HAEDAT records that were used for algorithm validations in this study (Supplementary Table 1), which also demonstrates the increase in sampling effort in the most recent years. Created using ArcMap 10.4.

Source Data

Extended Data Fig. 3 Performance of the CIE-fluorescence algorithm for phytoplankton bloom detection in 12 selected coastal oceans.

From left to right are the RGB-true color composite, ERGB composite, FLHRrc, and the bloom area (green pixels) detected by the CIE-fluorescence

algorithm. Created using ArcMap 10.4.

Source Data

Extended Data Fig. 4 Examples showing disadvantages of using NASA standard R_{rs} (i.e., with the removal of both Rayleigh and aerosol scattering) in algal bloom detection.

From left to right are the RGB composites, ERGB, nFLH, and the bloom areas (green pixels) detected by the CIE-fluorescence algorithm (based on R_{rc}, without the removal of aerosol scattering). Substantial amounts of invalid R_{rs} retrievals can be observed in the red-encircled areas in which severe blooms can be found. Additionally, nFLH shows high values at cloud edges (yellow-encircled areas), making it challenging to use a simple threshold to classify blooms. However, such problems can be circumvented in our CIE-fluorescence algorithm. Created using ArcMap 10.4.

Source Data

Extended Data Fig. 5 Sensitivity analysis of the impacts of aerosols on bloom detection.

(**a**) Responses of bloom area (BA) to changes in aerosol optical thickness (AOT). Aerosol reflectance (ρ_a) with AOTs of 0.01 and 0.02 at 869-nm is simulated and added to the MODIS images, and the resulting bloom areas (green pixels) with and without added ρ_a are compared. The left columns show the RGB composites, and the right three columns show the bloom areas under different AOTs. The percentages of BA changes are annotated in the panels. (**b**) The standard deviation between the 12 monthly mean values of AOT in global coastal waters (i.e., 66.7% of the intra-annual variability), and the histogram is shown in (**c**). Maps created using ArcMap 10.4.

Source Data

Extended Data Fig. 6 Comparison of different index-based algorithms in algal bloom detection in various coastal regions.

Image-specific thresholds (annotated within the panels) are required (labeled within the panels) for RI[50], ABI (estimated with FLH_{Rrc})[48], RBD[51], KBBI[51], and RDI[52] to delineate accurate bloom areas (i.e., high nFLH values, which appear as bright and darkish features on the ERGB images). The left panels are the bloom areas (green pixels) extracted using our CIE-fluorescence algorithm. The RGB-true color and ERGB composites are shown in Extended Data Fig. 3. Created using ArcMap 10.4.

Source Data

Extended Data Fig. 7 Annual median bloom count and the proportion of bloom-affected areas for large marine ecosystems (LMEs).

(**a**) Annual median bloom count, (**b**) proportion of bloom-affected areas. The data are ordered from the largest to the smallest. The LMEs are grouped by continent, and their names, numbers, and locations are shown in (a) and (b). Map created using Python 3.8.

Source Data

Extended Data Fig. 8 Comparison of bloom counts in the estuarine and non-estuarine regions.

Boxplots for long-term mean bloom count in the estuarine (n = 13,622 pixel observations) and non-estuarine (n = 361,604 pixel observations) regions. Comparison analysis was performed by two sided Welch's t-test ($P < 0.001$).Upper and lower bounds are first and third quartiles, the bar in the middle represents the median value, and the whiskers show the minimum and maximum values. Sixty-two estuarine zones from large rivers were selected, and the boundary of each zone was manually delineated according to high-resolution satellite images.

Source Data

Extended Data Fig. 9 Clusters of different bloom growth paths.

(a) The spatial distribution of different clusters. The fractions of different clusters across different latitudes are summarized. (**b**) The development of the maximum bloom-affected areas within a year within 1° × 1° grid cells, where all global grid cells are grouped into three distinct clusters according to the similarity of the bloom growth curve. The colored bond curves represent the mean values of all the grid cells, and their mean SST and associated standard deviations are shown with dashed lines and gray shading. The proportions of different clusters in the global bloom-affected areas are annotated. (**c**) and (**f**) The mean timing of the maximum bloom-affected areas (TMBAA) and the associated standard deviations between 2003 and 2019. The whole year in the Southern Hemisphere is shifted forward by 183 days in (**c**). Maps created using Python 3.8.

Source Data

Extended Data Fig. 10 Changes in climate extremes, global fertilizer uses, and fishery production over the past two decades.

(a) Changes in the bi-monthly Multivariate El Niño–Southern Oscillation (ENSO) index (MEI) between 2002 and 2020. Positive and negative MEI values represent EI Niño and La Niña events, respectively. The dots show annual mean values. **(b–c)** Trends of nitrogen and phosphorus from 2003 to 2019 for different countries. (**d**) Trends of fishery production from 2003 to 2018. Gray indicates no data. Maps created using ArcMap 10.4.

Source Data

Supplementary information

Supplementary Information

This file contains Supplementary Fig 1, and Supplementary Tables 1-3.

Reporting Summary

Peer Review File

Supplementary Data

This file contains source data for Supplementary Fig. 1.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 9

Source Data Extended Data Fig. 10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Dai, Y., Yang, S., Zhao, D. *et al.* Coastal phytoplankton blooms expand and intensify in the 21st century. *Nature* **615**, 280–284 (2023). https://doi.org/10.1038/s41586-023-05760-y

- Received: 17 May 2022
- Accepted: 25 January 2023
- Published: 01 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05760-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

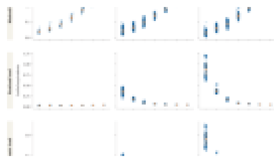
Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Coastal algal blooms have intensified over the past 20 years

Nature Research Briefing 01 Mar 2023

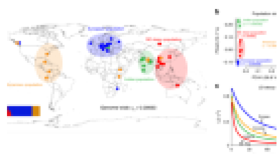
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05760-y



Genetic load: genomic estimates and applications in non-model animals

08 February 2022

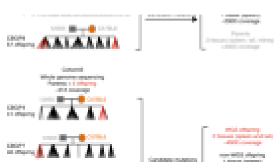
Giorgio Bertorelle, Francesca Raffini, … Cock van Oosterhout



Low genetic variation is associated with low mutation rate in the giant duckweed

18 March 2019

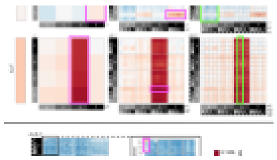
Shuqing Xu, Jessica Stapley, … Meret Huber



Similarities and differences in patterns of germline mutation between mice and humans

06 September 2019

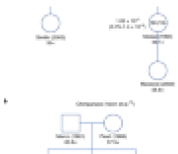
Sarah J. Lindsay, Raheleh Rahbari, … Matthew E. Hurles



[Extremely rare variants reveal patterns of germline mutation rate heterogeneity in humans]

14 September 2018

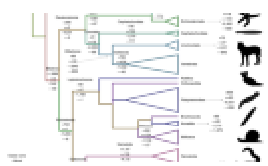
Jedidiah Carlson, Adam E. Locke, … The BRIDGES Consortium



[Direct estimation of mutations in great apes reconciles phylogenetic dating]

21 January 2019

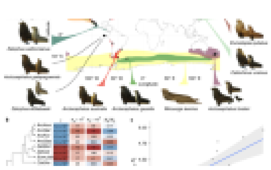
Søren Besenbacher, Christina Hvilsom, … Mikkel Heide Schierup



[Widespread patterns of gene loss in the evolution of the animal kingdom]

24 February 2020

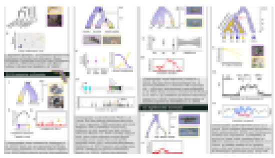
Cristina Guijarro-Clarke, Peter W. H. Holland & Jordi Paps



[Determinants of genetic variation across eco-evolutionary scales in pinnipeds]

08 June 2020

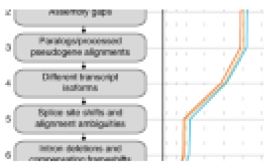
Claire R. Peart, Sergio Tusso, … Jochen B. W. Wolf



[The population genetics of crypsis in vertebrates: recent insights from mice, hares, and lizards]

09 August 2019

Rebecca B. Harris, Kristen Irwin, … Susanne P. Pfeifer



[A genomics approach reveals insights into the importance of gene losses for mammalian adaptations]

23 March 2018

Virag Sharma, Nikolai Hecker, … Michael Hiller

- Article
- Open Access
- [Published: 01 March 2023]

Evolution of the germline mutation rate across vertebrates

- Lucie A. Bergeron ,
- Søren Besenbacher,
- Jiao Zheng,
- Panyi Li,
- Mads Frost Bertelsen,
- Benoit Quintard,
- Joseph I. Hoffman,
- Zhipeng Li,
- Judy St. Leger,
- Changwei Shao,
- Josefin Stiller,
- M. Thomas P. Gilbert,
- Mikkel H. Schierup &
- …
- Guojie Zhang

Show authors

Nature **volume 615**, pages 285–291 (2023)

- 19k Accesses
- 259 Altmetric
- Metrics details

Abstract

The germline mutation rate determines the pace of genome evolution and is an evolving parameter itself[1]. However, little is known about what determines its evolution, as most studies of mutation rates have focused on single species with different methodologies[2]. Here we quantify germline

mutation rates across vertebrates by sequencing and comparing the high-coverage genomes of 151 parent–offspring trios from 68 species of mammals, fishes, birds and reptiles. We show that the per-generation mutation rate varies among species by a factor of 40, with mutation rates being higher for males than for females in mammals and birds, but not in reptiles and fishes. The generation time, age at maturity and species-level fecundity are the key life-history traits affecting this variation among species. Furthermore, species with higher long-term effective population sizes tend to have lower mutation rates per generation, providing support for the drift barrier hypothesis[3]. The exceptionally high yearly mutation rates of domesticated animals, which have been continually selected on fecundity traits including shorter generation times, further support the importance of generation time in the evolution of mutation rates. Overall, our comparative analysis of pedigree-based mutation rates provides ecological insights on the mutation rate evolution in vertebrates.

Main

Germline mutations are the proximate source of genomic innovation and inherited diseases[4]. Consequently, considerable effort has been spent on characterizing the molecular processes underlying these mutations and estimating germline mutation rates (GMRs). Mutations are rare events, yet the frequency at which they are introduced into genomes at each generation varies considerably across taxa, from approximately 10^{-11} mutations per site per generation in unicellular eukaryotes up to approximately 10^{-7} mutations per site per generation in multicellular eukaryotes[1,5,6]. Inferring the driving forces of GMR evolution has important implications for understanding the mechanisms underlying mutagenesis. Several hypotheses have been proposed to explain variation in GMRs among lineages. Some of these invoke molecular mechanisms such as DNA methylation[7] or microsatellite instability[8], whereas others invoke external factors such as exposure to mutagenic environments[9]. Other studies have argued that life-history traits might explain some of the variation both in the prevalence of mutations and in the ability to repair DNA. In particular, the generation time[10] and the metabolic rate[11] have been suggested to be key life-history traits that could be associated with germline mutations. From a long-term evolutionary

perspective, the 'drift barrier hypothesis' proposes that lower mutation rates may reflect the increased efficiency of natural selection at reducing the occurrence of mutations in species with large effective population sizes[3].

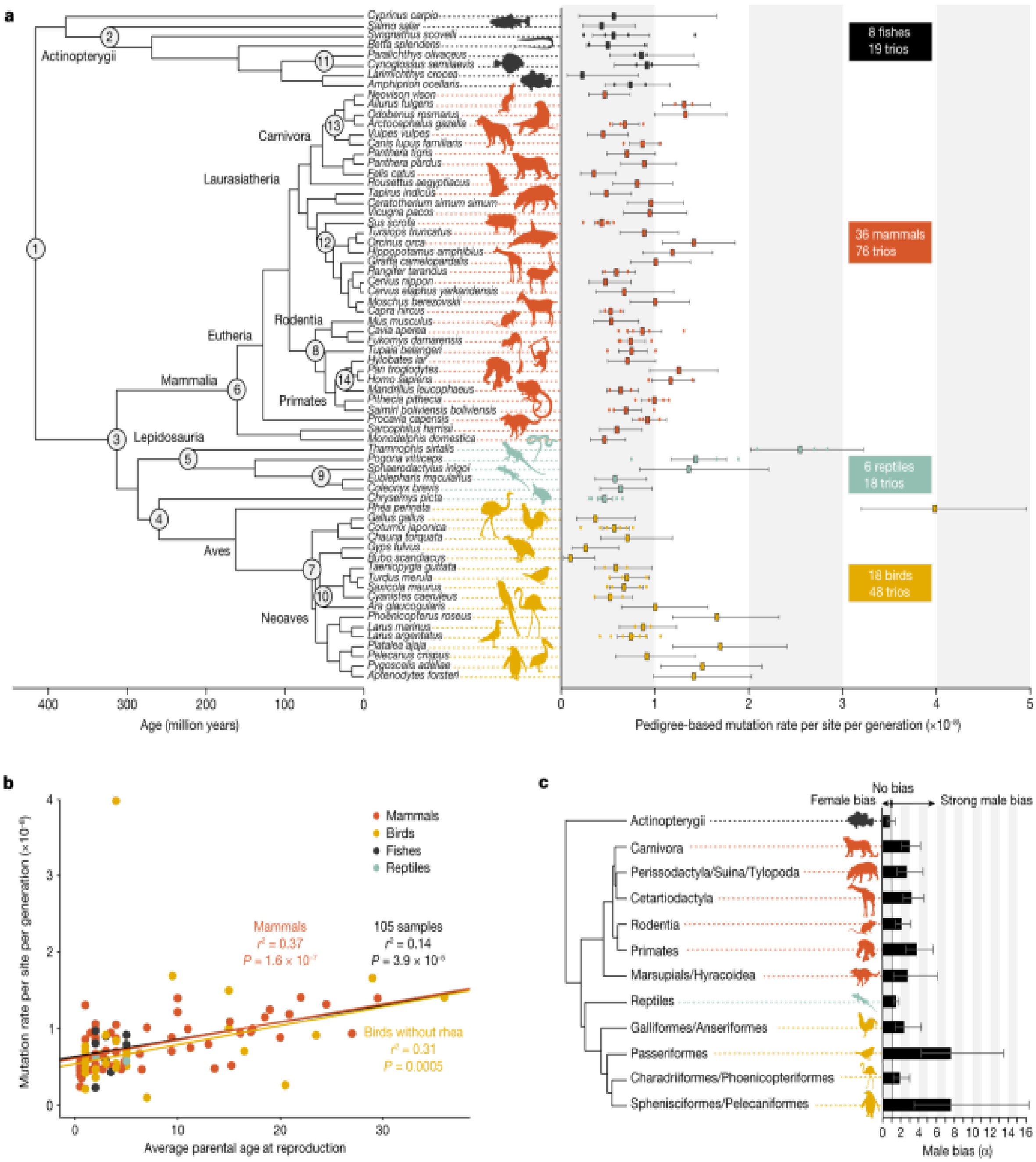
However, a lack of accurate and standardized GMR estimation has so far precluded testing current hypotheses of GMR evolution. Pedigree-based estimates of GMRs per generation have recently been published for a handful of vertebrate species, mainly focusing on humans and primates[12,13,14,15,16,17]. Furthermore, a recent comparative study of 16 mammalian species identified an effect of lifespan on somatic mutation rates inferred from the sequencing of intestinal crypts[18]. Nevertheless, interspecific comparisons of GMR variation remain restricted in taxonomic scope[19], partly due to the difficulty of comparing GMR estimates derived using different methodologies[2]. For example, alternative bioinformatic pipelines used in different studies can yield GMR estimates that vary by a factor of two, even when applied to the same parent–offspring trios[2]. This highlights the importance of applying consistent analytical pipelines for interspecies comparisons of GMRs. We therefore generated high-depth genome sequences (average coverage of more than 67×) for 323 individuals representing 151 trios of 68 vertebrate species, including 36 mammals, 18 birds, 8 ray-finned fishes and 6 reptiles (Supplementary Table 1). We then quantified species-specific GMRs across this wide range of vertebrate taxa using consistent bioinformatics pipelines to test long-standing evolutionary hypotheses on GMR evolution.

Per-generation mutation rate variation

We first estimated the per generation GMR ($\mu_{generation}$) for each trio (that is, mother, father and offspring) by comparing parental and offspring genomes (Fig. 1a, Supplementary Tables 2 and 3 and Supplementary Figs. 1–5 for details on the method). Overall, $\mu_{generation}$ varies by a factor of 40 across all species. On average, mutation rates per generation are higher in reptiles (average of all species 1.17×10^{-8}, 95% CI of the mean = 5.34×10^{-9} to 1.80×10^{-8}) and birds (average of all species 1.01×10^{-8}, 95% CI of the mean = 6.10×10^{-9} to 1.42×10^{-8}) than in mammals (average of all species 7.97×10^{-9}, 95% CI of the mean = 7.04×10^{-9} to 8.90×10^{-9}) and fishes

(average of all species 5.97×10^{-9}, 95% CI of the mean = 4.39×10^{-9} to 7.55×10^{-9}). However, the difference among the four major classes of vertebrates is not overall statistically significant (analysis of variance (ANOVA): $F = 1.86$, $P = 0.15$). Furthermore, the amount of variation in $\mu_{\text{generation}}$ among species tends to be higher for birds and lower for mammals and fishes (Fig. 1a), although this variation is arguably modest given large differences in life-history traits among these species (for example, there is a 2.8 million-fold difference in the body mass of killer whales and Siamese fighting fish, and there is a 93-fold difference in the generation time between humans and Texas banded geckos).

Fig. 1: Variation in GMRs and their association with life-history traits across 68 vertebrate species.



a, The phylogenetic tree of 68 species is based on UCE data and is calibrated with fossil data at 14 nodes (see Methods; Extended Data Fig. 3 and Supplementary Fig. 8). The average pedigree-based mutation rates per generation for each species, which are represented by the squares, show 40-fold variation among species. The 95% binomial confidence intervals are shown, and individual trios are represented by round points. See Supplementary Table 8 and Extended Data Fig. 4 for a comparison with published estimated rates of closely related species. **b**, The per-generation

mutation rate is significantly associated with the average parental age at the time of offspring production across all individuals with known paternal age (105 trios), using linear regression. For birds, this relationship is statistically significant after removing a single outlier, the Darwin's rhea. **c**, The male-to-female contribution ratio (α) is estimated for groups of vertebrates having at least 30 mutations phased to their parents of origin in each group. The highest male bias (7.6:1) is found in two bird lineages, whereas fishes and reptiles show negligible male bias. The data are represented with 95% confidence intervals based on the binomial variance. The silhouette of *Syngnathus scovelli* was created by J.S. All other silhouettes are from PhyloPic (http://phylopic.org), except for one of the silhouettes of *Sarcophilus harrisii*, which was created by S. Werning, and the silhouette of *Pan troglodytes*, which was created by T. M. Keesey (vectorization) and T. Hisgett (photography); both are available under a CC-BY 3.0 licence (https://creativecommons.org/licenses/by/3.0).

Species with longer generation intervals are expected to have higher per-generation mutation rates due to a combination of a larger number of cell divisions in spermatogenesis and more time for DNA damage to accumulate[12,13,14,20]. For the 105 trios for which parental age was known at reproduction, we found a significant positive association between $\mu_{generation}$ and the average parental age at reproduction (linear regression adjusted $r^2 = 0.14$, $P = 3.9 \times 10^{-5}$; Fig. 1b). This pattern is also significant for the 60 mammalian trios with known parental ages (linear regression adjusted $r^2 = 0.37$, $P = 1.6 \times 10^{-7}$) and for the 32 bird trios after excluding a single outlier, the Darwin's rhea (linear regression adjusted $r^2 = 0.31$, $P = 0.0005$). Furthermore, all three of these regressions have similar positive *y*-intercept values on the order of approximately 0.59×10^{-8} mutations per site per generation. For the trios with known parental ages, paternal and maternal ages at conception are strongly correlated (linear regression adjusted $r^2 = 0.77$, $P < 2.2 \times 10^{-16}$; Extended Data Fig. 1). However, multiple linear regression showed that the age of the father is the most significant explanatory variable (adjusted $r^2 = 0.15$, $P = 9.3 \times 10^{-5}$; paternal age $P = 0.018$; maternal age $P = 0.785$). Thus, a stronger effect of paternal than maternal age on the mutation rate seems to be universal for birds and mammals due to more germline mutations accumulating throughout the life of the male.

The specific types of de novo mutations (DNMs) observed across the 151 trios are concordant with the results of previous studies of individual species[12,13,14,21,22,23,24,25], including a ratio of transitions over transversions of 2.3 (95% CI on binomial distribution = 2.2–2.5) and a high proportion (48.5%, 95% CI on binomial distribution = 46.7–50.3%) of transitions from strong base pairing to weak base pairing (C:G > T:A) across all DNMs (Supplementary Table 4). Among C:G > T:A mutations, 42.4% (95% CI on binomial distribution = 39.9–45.0%) occurred at CpG sites. The direction of mutations from one base to another (that is, the spectrum of mutation) differed significantly across vertebrate classes ($\chi^2 = 30.0$, d.f. = 15, $P = 0.012$; Supplementary Table 4 and Supplementary Fig. 6). We also found significant differences among vertebrate classes for A > C mutations ($\chi^2 = 16.2$, d.f. = 3, $P = 0.001$) and for C > A mutations ($\chi^2 = 8.8$, d.f. = 3, $P = 0.032$). In particular, fish species exhibit significantly fewer A > C mutations and significantly more C > A mutations than the other vertebrate classes. However, this mutation pattern does not appear to be associated with genome-wide CG content, as overall, the CG content of fishes is similar to that of mammals and birds and lower than that of reptiles (Supplementary Fig. 7). Finally, there is no significant difference between the classes of species in the percentage of all mutations located in CpG sites ($\chi^2 = 4.3$, d.f. = 3, $P = 0.23$), implying that high mutation rates at CpG sites are a conserved feature across vertebrates.

Variable male-driven evolution

In mammals and birds, the much larger number of germ-cell divisions per generation in the male germ line leads to the expectation of a male mutation rate bias, coined the 'male-driven evolution hypothesis'[26,27]. However, very little is known about interspecific variation in the magnitude of the male-to-female ratio of the contribution of germline mutations (α). Previous studies have reported high α values in mammals (ranging from 1.0 to 20.1)[28] and birds (ranging from 3.9 to 6.5)[29] based on indirect estimates obtained by comparing rates of sequence divergence on the autosomes and sex chromosomes (see Extended Data Fig. 2 and Supplementary Table 5). However, other evolutionary forces can also act differently on the X chromosome and autosomes. For example, stronger natural selection on the

X chromosome could lead to lower than expected divergence from the common ancestor, upwardly biasing estimates of α[28]. Furthermore, estimates of α derived in this way are averages over a phylogenetic branch and may thus differ from the contemporary species α. Here we directly quantified α by assigning the parental origin of the DNMs. Around 48% of all 3,034 DNMs across all of the trios could be phased to their parental origin (see Supplementary Table 6 for positions of all mutations). Owing to the relatively small number of mutations in each trio (Supplementary Table 2), we analysed male bias after taxonomically grouping the species into classes and orders (Fig. 1c).

Mammals showed a male bias of α = 2.3 (95% CI = 2.0–2.6). In general, our α estimates are in line with previous estimates derived for similar species based on genome alignments[30,31]. For example, we found that among mammals, primates have the largest male bias with α = 3.8 (95% CI = 2.6–5.7), similar to what was previously reported for several species belonging to this group[12,13,14,21,22,32,33]. Rodents have the lowest male bias among the mammals in our study, with α = 2.1 (95% CI = 1.4–3.1), consistent with a previous study based on mouse pedigrees[34]. This pattern can be explained by the short generation time of rodents, which leads to a smaller difference in cell divisions between the male and female germ lines[35]. However, the variation in α is relatively small given the variation in generation time among species (for example, between 30 years for humans and 8 months for the short-tailed opossum). Thus, an alternative hypothesis to explain the observed α would be a higher contribution of DNA damage, specifically in the male germ line for species with large generation times[31].

Birds also showed an overall high male bias with α = 3.2 (95% CI = 2.5–4.1), although there is appreciable variation among different lineages. In particular, passerine birds and waterbirds (Pelecaniformes and Sphenisciformes) exhibited the largest male bias, both with α = 7.6 (95% CI = 4.3–13.5 for Passeriformes and 95% CI = 3.5–16.3 for Pelecaniformes and Sphenisciformes). High levels of male–male competition will lead to an increased amount of sperm being produced and faster sperm turnover, which would be expected to cause a higher male bias[36]. Indeed, many passerine birds have large cloacal protuberances[37] and relatively heavy testes[38], which are often used as proxies of sperm competition[39]. For instance, in two of the

passerine species included in our study, testes represent between 1.2% (for *Turdus merula*) and over 2% (for *Saxicola maurus*) of the total body mass[38]. Moreover, extra-pair mating is common in many passerine birds[40] as well as in penguins[41], also indicating a high level of sperm competition. Overall, our results lend further support to the male-driven hypothesis in birds and mammals[27].

By contrast, reptiles have a relatively small male bias with $\alpha = 1.5$ (95% CI = 1.2–1.8), whereas fishes appear to have a greater proportion of mutations of maternal origin ($\alpha = 0.8$), although the 95% CI overlaps 1 (95% CI = 0.5–1.4). This variation among vertebrate classes can be explained by differences in the process of gametogenesis. Although most birds and mammals produce sperm cells continuously through time[42], reptiles and fishes tend to be seasonal breeders, producing sperm cells during a limited period before the mating season[43,44,45], which will tend to reduce differences in cell division numbers between males and females, leading to more equal α. Moreover, female fishes are usually synchronous ovulators[46], producing hundreds to millions of eggs at the same time followed by a proliferation of new oogonia[47]. This implies that females continually produce germ cells throughout their life, which would further reduce the difference in cell division number between males and females.

Species with lower sex bias also exhibited a larger proportion of shared mutations between siblings, with 12.0% (s.e. of 6.5%) of shared mutations between siblings for fish and 8.1% (s.e. of 5.3%) for reptiles compared with 1.5% (s.e. of 0.7%) for mammals and 2.2% (s.e. of 1.4%) for birds (Supplementary Table 7). An explanation for the repeated occurrence of those mutations is that they appear during the primordial germ cell specification in one of the parents[48]. The occurrence of primordial germ cell specification mutations is independent of parental sex. Consequently, a higher number of primordial germ cell specification mutations in some vertebrate groups could be an alternative explanation for the lower male-biased contribution to DNMs.

Yearly mutation rates

To use our results for phylogenetic dating and to compare the speed of evolution among species with different generation times, we needed estimates of yearly mutation rates. Different methods have been used in the literature to estimate yearly mutation rates. When sample sizes are small, yearly rates are commonly inferred by dividing the per-generation rate by the average age of the parents (or the generation time if parental age is unknown)[49,50,51]. However, this method implicitly assumes a constant accumulation of mutations from conception to reproduction, that is, the regression line of mutation rate on parental age should run through the origin. Our results (Fig. 1b), as well as previous studies of mice, humans and cats[20,34], imply that parents always carry a minimum number of mutations in their gametes regardless of their age. This could lead to the yearly rate being overestimated for a given species if the sampled trio (or trios) had young parents compared with the average generation time for that species[52]. Consequently, we built a model that incorporates this mutational contribution at birth. Unfortunately, small per-species sample sizes in our dataset precluded modelling the effects of parental age separately for each species. However, we observed very similar intercepts and slopes across taxonomic groups, allowing us to fit a joint model for all species. A Poisson model explaining the number of mutations in each trio using a mutational contribution at birth and a weighted average of paternal and maternal age fits the data surprisingly well. To incorporate interspecific variation in male bias, we used the per-species fraction of paternal and maternal mutations estimated using read-backed phasing to weigh the average of the parental ages for each trio. Using this model, the number of predicted mutations matches the observed number with an overall r^2 of 0.73 (mammalian r^2 = 0.58, avian r^2 = 0.51; Supplementary Note 1).

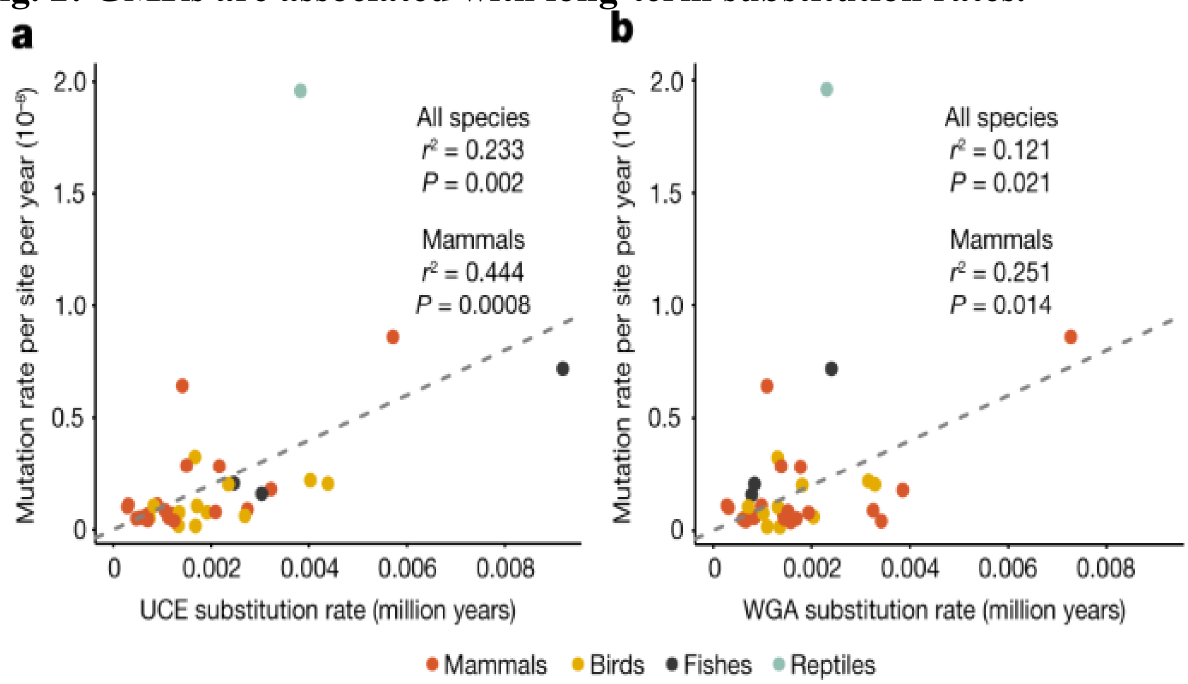
The yearly rates inferred with the naive method of dividing the per-generation rate by parental age (μ_{yearly}) and the rates obtained with our model ($\mu_{yearly_modelled}$) yielded similar results (Pearson's correlation r^2 = 0.40, P = 0.002), and for 55% of the species, μ_{yearly} falls within the 95% confidence interval of the $\mu_{yearly_modelled}$. As expected, the estimates showed the greatest differences for those species in which the parents reproduced far from the generation time, with the model-based estimates being smaller for those species that reproduced earlier than their generation time and larger for

those species that reproduced later than their generation time. For example, the pigs in our dataset reproduced at around 6 months of age, which is more than 5 years earlier than the estimated generation time of this species. Thus, $\mu_{yearly} = 8.64 \times 10^{-9}$ mutations per site per year was potentially overestimated compared with the $\mu_{yearly_modelled} = 1.05 \times 10^{-9}$ mutations per site per year at the generation time. Conversely, the yearly rate of the Texas banded gecko was potentially underestimated at $\mu_{yearly} = 3.17 \times 10^{-9}$ mutations per site per year using the reproductive age of 2 years of age from our dataset, whereas the modelled rate was $\mu_{yearly_modelled} = 1.96 \times 10^{-8}$ mutations per site per year at a generation time of between 3 and 4 months. Both the naive method and the modelled method have been used in the literature to estimate yearly rates and both have caveats owing to the underlying assumptions they require. Bearing this in mind, we decided to use $\mu_{yearly_modelled}$ for the current analysis as we believe that this measure is more representative of the yearly rate at the generation time for each species (estimated yearly rates are provided in Supplementary Table 9 for comparison).

The estimated average $\mu_{yearly_modelled}$ varies more than 120-fold among species (Supplementary Note 1 and Supplementary Table 9), with the highest $\mu_{yearly_modelled}$ estimated for the Texas banded gecko at 1.96×10^{-8} mutations per site per year (95% CI = 1.23×10^{-8} to 2.83×10^{-8}), whereas the lowest $\mu_{yearly_modelled}$ estimates were obtained for two bird species, the griffon vulture and the snowy owl, both with less than 0.18×10^{-9} mutations per site per year (snowy owl: $\mu_{yearly_modelled} = 0.16 \times 10^{-9}$, 95% CI = 0.05×10^{-9} to 0.34×10^{-9}; griffon vulture: $\mu_{yearly_modelled} = 0.17 \times 10^{-9}$, 95% CI = 0.07×10^{-9} to 0.32×10^{-9}). This large amount of interspecific variation is remarkable given that pedigree-based GMR estimates of individual species assessed by previous separate studies only show an approximately 16-fold variation in yearly GMRs[34,51]. Within primates, we observed a twofold variation across species and found a general trend for rates to be higher in the New World monkeys than in the great apes. This is consistent with previous independent estimates from primates[19] and supports the ‘hominoid slowdown’ hypothesis[53,54,55,56].

Next, we used $\mu_{yearly_modelled}$ to assess the strength of the association between GMRs and long-term evolutionary substitution rates. To obtain an estimate of the long-term substitution rate, we used the alignment of ultraconserved elements (UCEs), which are more likely to align among taxonomically distant species, plus 1,000 bp of flanking regions on each side of the UCE sequences, which will more closely reflect the neutral substitution rate[57]. We found a significant positive correlation between $\mu_{yearly_modelled}$ and the UCE substitution rate after excluding domesticated species owing to their overall much higher yearly mutation rates (see the following section; phylogenetic generalized least squares (PGLS): adjusted $r^2 = 0.23$, $P = 0.002$; Fig. 2a). This pattern is especially pronounced for mammals (PGLS: adjusted $r^2 = 0.44$, $P = 0.0008$), even after removing the two outliers (PGLS: adjusted $r^2 = 0.32$, $P = 0.009$). We also found a significant relationship between $\mu_{yearly_modelled}$ and the long-term substitution rate inferred using whole-genome alignments (PGLS: adjusted $r^2 = 0.12$, $P = 0.02$; Fig. 2b).

Fig. 2: GMRs are associated with long-term substitution rates.



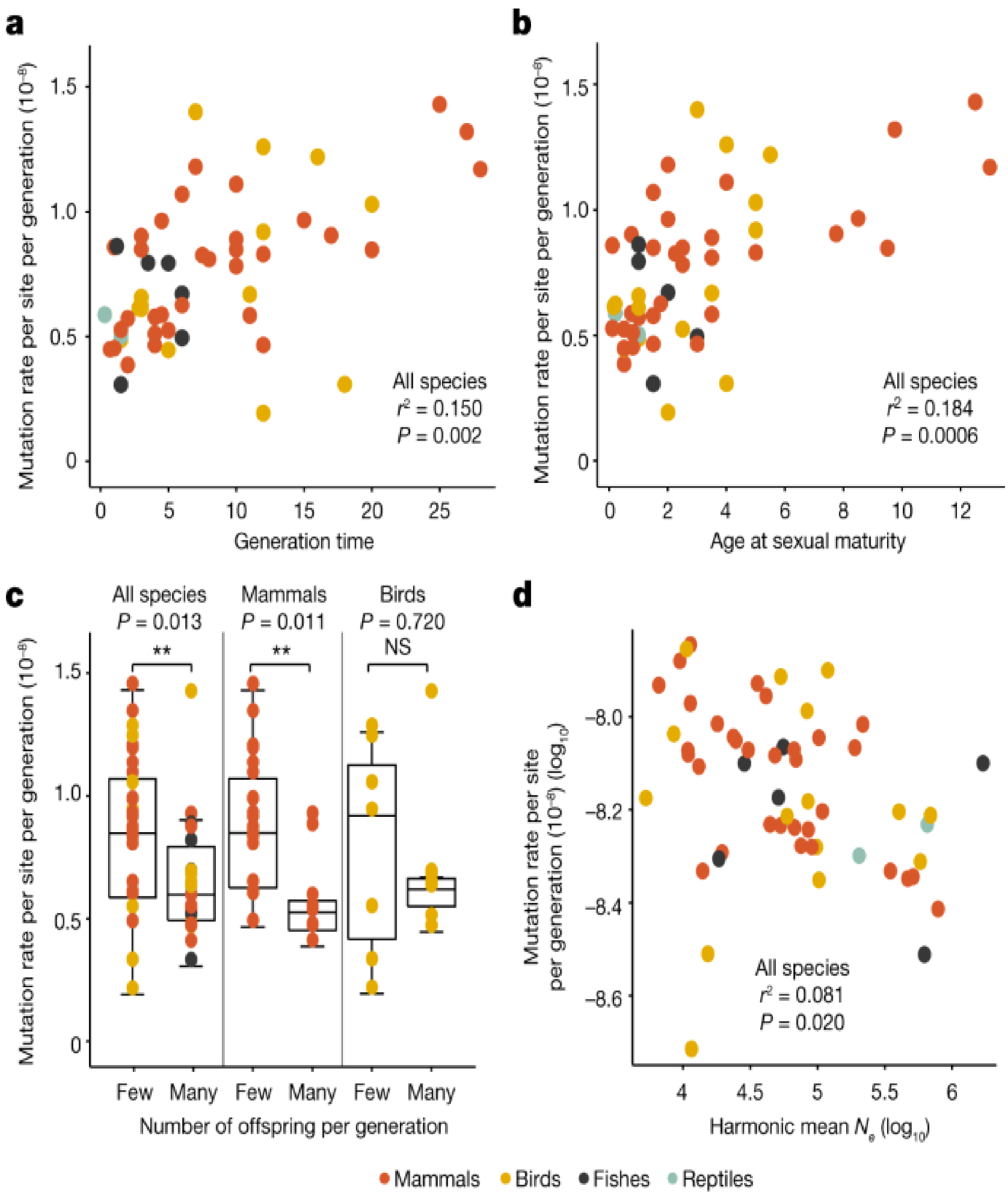
a,**b**, There is a positive association between the modelled yearly pedigree-based mutation rates and the macroevolutionary substitution rates when using phylogenetic regression (PGLS) on both UCEs and their flanking

sequences (**a**) and whole-genome alignments (WGAs) (**b**). The grey dashed lines indicate equality. See Extended Data Fig. 5 for plots of the same data on a log scale and Extended Data Fig. 6 for a comparison of UCE and WGA methods.

Life-history traits shape GMR variation

To test various hypotheses relating to the causes of GMR variation among species, we tested for associations between the modelled mutation rate per generation ($\mu_{generation_modelled}$) and life-history traits including mating system (monogamy versus polygamy), maturation time, body mass, longevity, fecundity and the generation time (Supplementary Table 9). We used the $\mu_{generation_modelled}$ instead of the $\mu_{generation}$ as the former is less dependent on the age of the parents and is more representative of the rate at generation time for a given species. Although taking into account phylogenetic relatedness, many of these traits are significantly associated with $\mu_{generation_modelled}$ including the generation time (PGLS: adjusted $r^2 = 0.15$, $P = 0.002$; Fig. 3a), the maturation time (PGLS: adjusted $r^2 = 0.18$, $P = 0.0006$; Fig. 3b) and the number of offspring per generation (PGLS: adjusted $r^2 = 0.10$, $P = 0.013$; Fig. 3c). Species with a higher number of offspring per generation also showed significantly lower $\mu_{generation_modelled}$ when considering only mammalian species (PGLS: adjusted $r^2 = 0.17$, $P = 0.011$), but this relationship was not significant for birds (PGLS: adjusted $r^2 = -0.066$, $P = 0.720$). Collectively, these traits explain almost 18% of the variation in $\mu_{generation_modelled}$ (multiple PGLS: adjusted $r^2 = 0.18$, $P = 0.004$). The other life-history traits that we tested, including longevity, mating strategy and body mass, are not significantly associated with $\mu_{generation_modelled}$ (see Extended Data Fig. 7).

Fig. 3: Predictors of interspecific variation in GMRs.



a–**c**, Significant positive associations are found using phylogenetic regression (PGLS) between the modelled per-generation mutation rates and three life-history traits: species-specific mean generation time (**a**), age at sexual maturity (**b**) and the number of offspring per generation (**c**). In total there are 55 species with modelled per-generation rates, including 32 mammalian and 15 avian species. The box plot in **c** represents the median,

the interquartile range and the maximum and minimum excluding outliers. **d**, Species-specific average per-generation mutation rates are negatively associated with the harmonic mean of the effective population size (N_e) over the past 1 million years, using phylogenetic regression (PGLS).

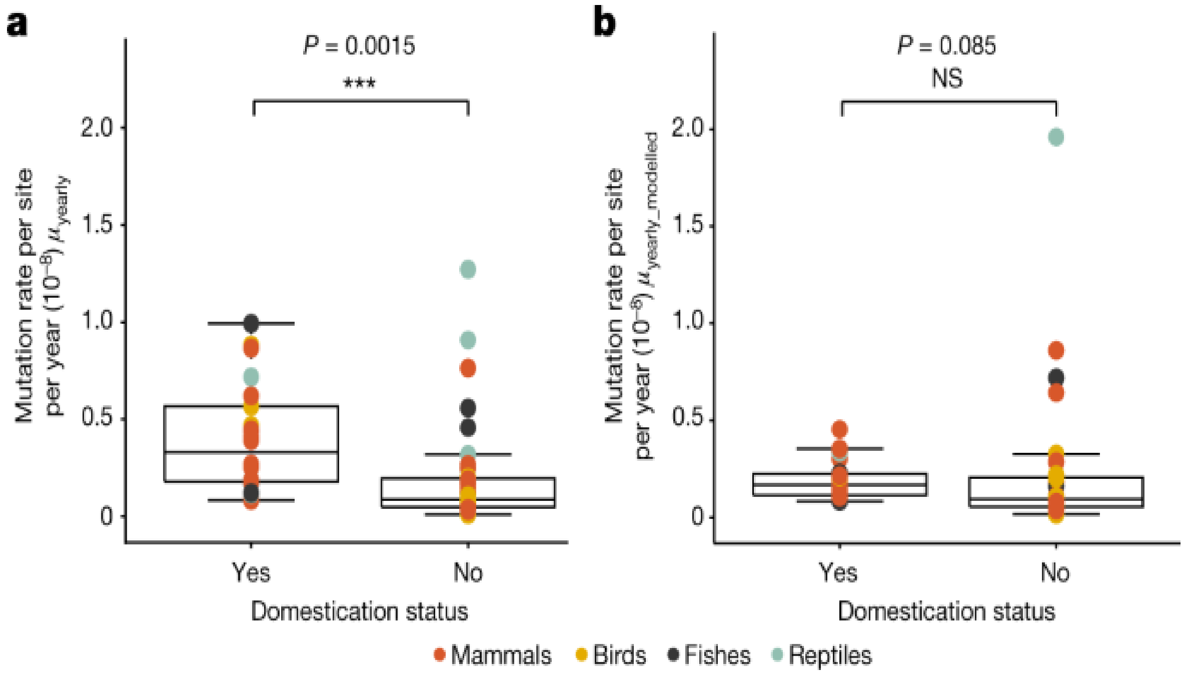
Another key parameter for species evolution is the effective population size (N_e), which impacts genetic drift and the efficacy of selection. To investigate the effect of N_e on $\mu_{generation_modelled}$ and to test the drift barrier hypothesis[3], which predicts the evolution of higher mutation rates in species with small N_e, we calculated N_e using the pairwise sequentially Markovian coalescent method based on one randomly selected father per species. To avoid circularity, we estimated N_e based on the substitution rate calculated from the UCE alignment (Supplementary Table 9). Indeed, if N_e was estimated using the pedigree-based mutation rate, a stronger correlation might arise between N_e and the mutation rate (see Extended Data Fig. 8). We found a significant negative association between $\mu_{generation_modelled}$ and the harmonic mean N_e per species over the past 30,000–1,000,000 years (PGLS: adjusted $r^2 = 0.08$, $P = 0.020$; Fig. 3d) as would be expected under the drift barrier hypothesis. This relationship is mainly driven by mammals (PGLS: adjusted $r^2 = 0.31$, $P = 0.0006$), a signal that is also observed when using the harmonic average N_e over a smaller timescale (30,000–130,000 years; PGLS: adjusted $r^2 = 0.10$, $P = 0.04$, Extended Data Fig. 8). The most appropriate timeframe used to estimate N_e depends on the evolutionary time necessary for the mutation rate to adapt to changes in N_e. However, the pairwise sequentially Markovian coalescent method cannot accurately estimate recent N_e. To overcome this limitation, we also estimated N_e as $\pi/4\mu$, with nucleotide diversity (π) and the substitution rate per site per generation (μ) estimated from the UCE alignments. This results in a similar negative association between N_e and $\mu_{generation_modelled}$ (linear regression: adjusted $r^2 = 0.83$, $P = 2.2 \times 10^{-16}$; Extended Data Fig. 9), further supporting the drift barrier hypothesis. However, caution should be taken as N_e estimates rely on generation times inferred from contemporary observations, whereas generation times could conceivably have changed over evolutionary timescales. Furthermore, population size depends negatively on the

generation time (PGLS N_e in log scale: adjusted $r^2 = 0.20$, $P = 0.0004$). Therefore, a negative association between N_e and μ could potentially be driven by a large effect of the generation time on per-generation mutation rates.

High yearly rates in domesticated species

Domestication imposes strong artificial selection, recurrent genetic bottlenecks or both. Our dataset includes 22 domesticated or semi-wild species that have been bred in captivity for many generations. When using the naive method of dividing the per-generation rate by the parental age, these species show significantly higher μ_{yearly} than the non-domesticated species (PGLS: adjusted $r^2 = 0.13$, $P = 0.0015$; Fig. 4a). The higher mutation rates of domesticated animals are likely due to strong artificial selection for traits such as shorter generation times. Indeed, using $\mu_{yearly_modelled}$, we found no difference between domesticated and non-domesticated species (PGLS: adjusted $r^2 = 0.037$, $P = 0.08$; Fig. 4b). Consequently, the higher yearly mutation rate observed in domesticated species is more likely to be explained by the lowering of reproductive age associated with domestication rather than by an inherent change to the mutational process caused by relaxed selection on the mutation rate due to small population sizes and bottlenecks associated with domestication[58,59].

Fig. 4: The yearly GMRs are higher in domesticated species than in non-domesticated species.



a, Yearly GMRs are significantly higher in domesticated or farmed species than in wild species (using phylogenetic regression (PGLS) on a total of 68 species). **b**, Using the modelled mutation rate instead (using phylogenetic regression (PGLS) on a total number of 55 species) shows that there is no difference in yearly GMRs between domesticated and non-domesticated animals, suggesting that this difference is mainly driven by the shorter generation time of domesticated species. The box plots represent the median, the interquartile range and the maximum and minimum excluding outliers.

Conclusions

Here we analysed pedigree-based GMR variation in an unprecedentedly broad phylogenetic context. We showed that there is a consistent male bias in mammals and birds, whereas reptiles and fish exhibited more evenly matched contributions of DNMs between parents. This could be due to contrasting mutagenic processes, such as differences in male and female germline cell division observed in mammals and birds, or differences among species in the proportion of DNMs occurring in primordial germ cell specification versus in the parental germ lines. Our results also support the drift barrier hypothesis, as we found a negative association between the per-generation mutation rate and effective population size. Moreover, our results

suggest that an appreciable proportion of the variation in the GMR can be explained by life-history traits, including maturation time and the number of offspring per generation. Our study also highlights the importance of the generation time, as illustrated by the particular case of domesticated animals, in which exceptionally high yearly mutation rate estimates can be explained by artificially induced short generation times. In addition, some of the trio samples in our study were collected from captive animals at zoos or conservation centres. These populations might have different generation times than those in the wild, which could potentially introduce biases into some of our mutation rate estimates. Future studies should focus on wild pedigree samples, which can be accessed from long-term conservation and monitoring programmes[60].

Methods

Samples

Samples were collected from zoos, zoological museums, research institutes and farms from all over the world. Samples were provided from collaborators for research that was undertaken at the Natural History Museum of Denmark, permit 2020-12-7186-00733 from the Danish Ministry of Environment and Food, and when applicable, CITES Certificate of Scientific Exchange number DK003. Genomic DNA was extracted using DNeasy Blood and Tissue Kits (Qiagen) following the manufacturer’s instructions. BGIseq libraries were built in China National GeneBank (CNGB), Shenzhen, China, and whole-genome paired-end sequencing (read length 2 × 100 bp) were performed on the BGISEQ500 platform. We aimed for 60–80× raw sequence coverage per sample. A total of 68 species for which a reference genome was available were retained in the final dataset, representing 151 trios for which whole blood or other tissue material was available for DNA extraction and for which parentage had been genetically determined[61]. Information on the samples is provided in Supplementary Table 1.

GMR estimation

We applied a similar bioinformatic analysis pipeline to our previous study of rhesus macaques[12]. Raw reads were trimmed with SOAPnuke filter[62]. The mapping was conducted with BWA-MEM version 0.7.15 (ref. [63]). The versions of the reference genomes for each species are provided in Supplementary Table 9. A post-mapping step removed any reads mapping to multiple regions of the genome as well as duplicated reads using Picard MarkDuplicates 2.7.1. We called variants for each individual using HaplotypeCaller in BP-RESOLUTION mode with GATK 4.0.7.0 (ref. [64]). This mode returns a genotype quality and depth for all positions of the genome, not only the polymorphic sites. As recommended by GATK best practices, GenomicsDBImport combined all gVCF files per species into a single file and GenotypeGVCF applied a joint genotyping of all samples within a given species (see Supplementary Table 3 with details of raw sequences coverage, mapping quality, and coverage after mapping and variant calling). Similar filtering methods to those in our previous study were then applied to detect DNMs[12]. Therefore, each trio was filtered as followed:

1. (1)

   For site filtering, the variant positions were filtered with the following parameters: QualByDepth (QD) < 2.0, FisherStrand (FS) > 20.0, RMSMappingQuality (MQ) < 40.0, MQRankSum < −2.0, MQRankSum > 4.0, ReadPosRankSum < −3.0, ReadPosRankSum > 3.0 and StrandOddsRatio (SOR) > 3.0 according to previously tested filters[12].

2. (2)

   For Mendelian violations, variants that deviated from Mendelian inheritance were selected using GATK SelectVariant and refined with an R script to keep only sites in which both parents were homozygous for the reference allele (HomRef), and the offspring was heterozygous (Het).

3. (3)

For allelic balance filter, in the case of a DNM, approximately 50% of the reads in the offspring should support the alternative alleles. Our allelic balance filter cut-off was 30–70% of the reads supporting the alternative allele, similar to previous studies[12,32,65,66].

4. (4)

   For depth filter (DP), only positions with a DP > $0.5 \times m_{depth}$ and DP < $2 \times m_{depth}$ for each individual were kept, with m_{depth} being the average depth of the trio. This strict DP filter minimized the effects of sequencing errors in regions of low sequencing depth and mis-mapping errors in high-coverage regions.

5. (5)

   For genotype quality filter (GQ), to ensure that only high-quality genotypes were retained for the analysis of trios, we removed all sites where one individual of the trio had a GQ < 60 (see Supplementary Fig. 2 for a comparison of various GQ thresholds on a subset of species).

In addition, we called variants with bcftools (version 1.2)[67] in the region of the candidate DNMs and removed the sites that appeared as false-positive calls (that is, at least one parent had the same variant as the offspring or the offspring had no variant). The number of candidates discarded varied among species (Supplementary Table 2). This quality control step produced similar results to a manual check with IGV[68]. Moreover, calling variants with different variant callers has been shown to be an efficient method to reduce false-positive calls[2]. All positions of DNMs are provided in Supplementary Table 6. In addition, we showed that sample type, reference genome quality and mapping quality can affect the results on the number of candidates, the false-positive rate and false-negative rate (FNR), yet, the estimated mutation rates are not affected (Supplementary Figs. 3–5).

To estimate per-generation rates, we divided the number of candidate DNMs, without the apparent false-positive candidates, per the callable genome. A site was considered callable when it passed the same filters as the polymorphic sites, that is, when both parents were HomRef (filter 2) and the three individuals passed the depth filter (filter 4) and the genotype quality

threshold (filter 5). On the sites considered callable, we applied a correction for the FNR, that is, the proportion of sites where true DNMs will not be called as such. Two methods have been used in the literature to estimate FNR: one is the simulation of mutations and the other is a correction on the filters that are not accounted for in the callable genome. As in our previous study of GMR[12], we used the latter method, which is more conservative. This corrected for the remaining filters that can only be applied on polymorphic sites, such as the site filters and the allelic balance filter (filter 2). We estimated the proportion of sites that would be filtered away by the site filters on the parameters following a known distribution (FS, MQRankSum and ReadPosRankSum), and the expected sites filtered away by the allelic balance filter as the number of true heterozygote sites (one parent HomRef, the other parent HomAlt and their offspring Het) outside the allelic balance threshold. The mutation rate per site per generation was then estimated per trio as $\mu_{\text{generation}} = \text{DNMs}/((1 - \text{FNR}) \times 2 \times \text{CG})$. We estimated the 95% binomial confidence interval per species using the binconf() function in R, with the default Wilson score.

To calculate yearly rates (μ_{yearly}), we divided the per-generation rate by the average age of the parents at the time of reproduction weighted by the relative contribution of each parent (inferred with α for 105 trios) or by the generation time (for 46 trios without parental ages). The resulting μ_{yearly} estimates were averaged per species (for 29 species with multiple trios available). These yearly rates are dependent on the age of reproduction of the parents. Therefore, to calculate a yearly rate at generation time, we first modelled how the mutation rate of a trio was affected by the weighted average of the parental ages (using the paternal fraction estimated for that species as a weight). We then extended the model to fit how each species deviated from the average and used this to correct for differences between the observed reproductive age in our dataset and the expected generation time of a species (see Supplementary Note 1). With this, we estimated a new $\mu_{\text{yearly_modelled}}$ and a $\mu_{\text{generation_modelled}}$ that are more representative of the rate at generation time for each species.

Phylogenetic analysis

The phylogeny was built based on two sets of UCEs: 5,472 baits for 5,060 UCEs in tetrapods[57] and 2,628 baits for 1,314 UCEs in acanthomorphs[69]. We used the Phyluce software[70] to locate the probes in the reference genomes of our 68 species with 6 additional species contained in our original dataset. We extracted a flanking region of ±1,000 bp for each probe and aligned them with Mafft aligner version 7.470 (ref. [71]). We then created a 75% completion matrix, that is, each alignment contains at least 75% of the taxa (55 species), resulting in 63 alignments from the acanthomorph set and 2,742 probes from the tetrapod set (all alignments are available on Figshare). A phylogenetic tree was built using IQ-TREE version 2.0.3 (ref. [72]), with the appropriate substitution model inferred for each of the 2,805 alignments, a maximum likelihood tree search and 1,000 bootstrap replicates. To validate our tree, we also estimated a second tree based on a MultiZ alignment to the human genome and obtained similar results (Extended Data Fig. 9). The phylogenetic tree was calibrated to absolute time using the chronos function of the 'ape' package in R, with a smoothing parameter lambda of 0 and a 'relaxed' model[73,74]. Fourteen nodes were calibrated following previously published calibrations. The robustness of the tree was assessed by removing each node independently (see Extended Data Fig. 3).

1. (1)

   Actinopterygii/Sarcopterygii: divergence time 416 million years ago (Ma), upper bound 425.4 Ma[75]

2. (2)

   The first node in the Actinopterygii group: divergence time 378.2 Ma[76]

3. (3)

   Sauropsida (birds and reptiles)/Synapsida (mammals): divergence time 313.4 Ma[77]

4. (4)

   Archosauria (birds)/Testudines: divergence time 260 Ma[78]

5. (5)

   The basal nodes of the Lepidosauria: divergence time 222.8 Ma[79]

6. (6)

   First mammalian node, Eutheria/Metatheria: divergence time 160.7 Ma[75]

7. (7)

   Galloanserae/Neoaves: divergence time 66 Ma[77]

8. (8)

   Glire/Primates: divergence time 61.7 Ma[77]

9. (9)

   Basal gekkotan node: divergence time 54 Ma[80]

10. (10)

    Passeriformes/Psittaciformes: divergence time 51.81 Ma[81]

11. (11)

    Cynoglossidae/Paralichthyidae: divergence time 50 Ma[76]

12. (12)

    Sus scrofa/other Cetartiodactyla: divergence time 48.5 Ma[77]

13. (13)

    Canidae/Arctoidea: divergence time 37.1 Ma[75]

14. (14)

Hominoidea/Cercopithecoidea: divergence time 23.5 Ma[77]

Mutational spectrum and sex bias

To analyse the spectrum of mutation, we grouped the trios into higher taxonomic levels, that is, mammals, birds, fishes and reptiles. Thus, the percentages reported are based on the total candidate mutations from each group of species. We explored the genomic context of the mutations from a C or a G base to determine whether they were located in CpG sites (respectively followed by a G or preceded by a C) (see Supplementary Table 4). We phased the DNMs to their parental origin using the read-backed phasing method described previously (GitHub: https://github.com/besenbacher/POOHA)[82]. This method uses the read-pairs containing a DNM and another heterozygous variant to determine the parental origin of the mutation when the heterozygous variant is present in both the offspring and one of the parents. The phasing allowed us to identify parental biases in the contribution of the DNMs by grouping multiple species to increase the number of phased mutations and obtain a minimum of 30 phased mutations per taxon. From this analysis, we omitted the Egyptian roussette (*Rousettus aegyptiacus*), Chinese tree shrew (*Tupaia belangeri*), griffon vulture (*Gyps fulvus*), blue-throated macaw (*Ara glaucogularis*), snowy owl (*Bubo scandiacus*) and Darwin's rhea (*Rhea pennata*), as these could not be grouped with another monophyletic clade. To quantify the effect of parental age, a linear regression between the per-generation mutation rate and the average parental age at the time of reproduction was implemented using the lm function in R. Multiple linear regression was also used to identify whether paternal or maternal age was the strongest predictor of the empirical mutation rate.

Life-history trait analysis

We tested the effect of various life-history traits (fitted as continuous and discrete variables) on the yearly rate for each species using PGLS analysis in the R package 'caper'[83] (see Supplementary Table 9 for details about each life-history trait).

Effective population size

We used pairwise sequentially Markovian coalescent (PSMC) models to estimate the effective population size of each species[84]. Fastq sequences were obtained using bam format aligned sequences of one randomly selected father per species and were converted into fastq format using samtools mpileup command and vcf2fq. As recommended, the minimum depth was set to one-third of the average for the sample and twice the average for the maximum. For mammals, fish and reptiles, the parameters of the PSMC were set to –N25 for the maximum number of iterations of the algorithm, –t15 as the upper limit for the time to the most recent common ancestor, –r5 for the initial θ/ρ value, and finally the atomic intervals –p of '4 + 25 × 2 + 4 + 6'. These parameters were used previously for PSMC analysis of various species, including primates[84,85], cetaceans[86], Felidae[87], fishes[88,89] and turtles[90]. For birds, we used different parameters according to the literature with –N30 –t5 –r5 (ref. [91]). Finally, to simulate the history inferred by PSMC, we parameterized the generation time and the mutation rate inferred from the UCE alignment. We then explored the effect of the harmonic mean N_e over windows of 30,000 years to 1,000,000 years. We also compared N_e estimated obtained with this method with those estimated based on $N_e = \pi/4\mu$. Nucleotide diversity (π) was calculated using ANGSD[92]. This approach was implemented in three consecutive steps. From the alignment files, a global estimate of the site frequency spectrum was inferred using a maximum likelihood method, then the empirical π value was estimated per site, and finally, a sliding window approach was used to estimate π for each species. We used a window size of 50 kb and a step size of 10 kb together with an average pairwise estimation of the π to obtain global estimates of π. This analysis was restricted to unrelated individuals from each species, which corresponded to the 2 unrelated parents for 55 species, between 3 and 7 individuals for 10 species, and 3 species were excluded from this analysis as the parents were first-degree relatives.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Whole-genome sequences of all species except humans are accessible in the National Center for Biotechnology Information under the BioProject ID PRJNA767781. The human sequences are available on request to L.A.B. and should be used only for GMR studies, based on the participant's request. The alignments for the UCE tree are available on Figshare (https://doi.org/10.6084/m9.figshare.19221693.v1). All animal silhouettes are from PhyloPic (http://phylopic.org/), except for the silhouette of *S. scovelli*, which was created by J.S. The silhouette of *P. troglodytes* was created by T. M. Keesey (vectorization) and T. Hisgett (photography), and the one of *S. harrissi* silhouettes was created by S. Werning; both are available under a CC-BY 3.0 license (https://creativecommons.org/licenses/by/3.0/); the other silhouettes are available under a Public Domain Mark 1.0 licence.

Code availability

The bioinformatics pipeline to analyse the genomes and all other data analyses are available on GitHub (https://github.com/lucieabergeron/vertebrate_rate).

References

1. Lynch, M. et al. Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.* **17**, 704–714 (2016).

2. Bergeron, L. A. et al. The mutationathon highlights the importance of reaching standardization in estimates of pedigree-based germline mutation rates. *eLife* **11**, e73577 (2022).

3. Lynch, M. Evolution of the mutation rate. *Trends Genet.* **26**, 345–352 (2010).

4. Acuna-Hidalgo, R., Veltman, J. A. & Hoischen, A. New insights into the generation and role of de novo mutations in health and disease. *Genome Biol.* **17**, 241 (2016).

5. Sturtevant, A. H. Essays on evolution. I. On the effects of selection on mutation rate. *Q. Rev. Biol.* **12**, 464–467 (1937).

6. Zhang, G. The mutation rate as an evolving trait. *Nat. Rev. Genet.* **24**, 3 (2022).

7. Mugal, C. F., Arndt, P. F., Holm, L. & Ellegren, H. Evolutionary consequences of DNA methylation on the GC content in vertebrate genomes. *G3* **5**, 441–447 (2015).

8. Baer, C. F., Miyamoto, M. M. & Denver, D. R. Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nat. Rev. Genet.* **8**, 619–631 (2007).

9. Wright, S. D., Ross, H. A., Jeanette Keeling, D., McBride, P. & Gillman, L. N. Thermal energy and the rate of genetic evolution in marine fishes. *Evol. Ecol.* **25**, 525–530 (2011).

10. Ohta, T. An examination of the generation-time effect on molecular evolution. *Proc. Natl Acad. Sci. USA* **90**, 10676–10680 (1993).

11. Martin, A. P. & Palumbi, S. R. Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl Acad. Sci. USA* **90**, 4087–4091 (1993).

12. Bergeron, L. A. et al. The germline mutational process in rhesus macaque and its implications for phylogenetic dating. *Gigascience* **10**, giab029 (2021).

13. Wu, F. L. et al. A comparison of humans and baboons suggests germline mutation rates do not track cell divisions. *PLoS Biol.* **18**, e3000838 (2020).

14. Wang, R. J. et al. Paternal age in rhesus macaques is positively associated with germline mutation accumulation but not with measures of offspring sociability. *Genome Res.* **30**, 826–834 (2020).

15. Campbell, C. R. et al. Pedigree-based and phylogenetic methods support surprising patterns of mutation rate and spectrum in the gray

mouse lemur. *Heredity* **127**, 233–244 (2021).

16. Besenbacher, S., Hvilsom, C., Marques-Bonet, T., Mailund, T. & Schierup, M. H. Direct estimation of mutations in great apes reconciles phylogenetic dating. *Nat. Ecol. Evol.* **3**, 286–292 (2019).

17. Thomas, G. W. C. et al. Reproductive longevity predicts mutation rates in primates. *Curr. Biol.* **28**, 3193–3197.e5 (2018).

18. Cagan, A. et al. Somatic mutation rates scale with lifespan across mammals. *Nature* **604**, 517–524 (2022).

19. Chintalapati, M. & Moorjani, P. Evolution of the mutation rate across primates. *Curr. Opin. Genet. Dev.* **62**, 58–64 (2020).

20. Wang, R. J. et al. De novo mutations in domestic cat are consistent with an effect of reproductive longevity on both the rate and spectrum of mutations. *Mol. Biol. Evol.* **39**, msac147 (2022).

21. Venn, O. et al. Strong male bias drives germline mutation in chimpanzees. *Science* **344**, 1272–1275 (2014).

22. Jónsson, H. et al. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).

23. Tatsumoto, S. et al. Direct estimation of de novo mutation rates in a chimpanzee parent-offspring trio by ultra-deep whole genome sequencing. *Sci. Rep.* **7**, 13561 (2017).

24. Yuen, R. K. C. et al. Genome-wide characteristics of de novo mutations in autism. *npj Genomic Med.* **1**, 160271–1602710 (2016).

25. Wang, H. & Zhu, X. De novo mutations discovered in 8 Mexican American families through whole genome sequencing. *BMC Proc.* **8**, S24 (2014).

26. Li, W.-H., Yi, S. & Makova, K. Male-driven evolution. *Curr. Opin. Genet. Dev.* **12**, 650–656 (2002).

27. Miyata, T., Hayashida, H., Kuma, K., Mitsuyasu, K. & Yasunaga, T. Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 863–867 (1987).

28. Wilson Sayres, M. A. & Makova, K. D. Genome analyses substantiate male mutation bias in many species. *BioEssays* **33**, 938–945 (2011).

29. Ellegren, H. & Fridolfsson, A.-K. Male-driven evolution of DNA sequences in birds. *Nat. Genet.* **17**, 182–184 (1997).

30. Sayres, M. A. W., Venditti, C., Pagel, M. & Makova, K. D. Do variations in substitution rates and male mutations bias correlate with life-history traits? A study of 32 mammalian genomes. *Evolution* **65**, 2800–2815 (2011).

31. de Manuel, M., Wu, F. L. & Przeworski, M. A paternal bias in germline mutation is widespread in amniotes and can arise independently of cell division numbers. *eLife* **11**, e80008 (2022).

32. Francioli, L. C. et al. Genome-wide patterns and properties of de novo mutations in humans. *Nat. Genet.* **47**, 822–826 (2015).

33. Gao, Z. et al. Overlooked roles of DNA damage and maternal age in generating human germline mutations. *Proc. Natl Acad. Sci. USA* **116**, 9491–9500 (2019).

34. Lindsay, S. J., Rahbari, R., Kaplanis, J., Keane, T. & Hurles, M. E. Similarities and differences in patterns of germline mutation between mice and humans. *Nat. Commun.* **10**, 4053 (2019).

35. Gibbs, R. A. et al. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493–520 (2004).

36. Blumenstiel, J. P. Sperm competition can drive a male-biased mutation rate. *J. Theor. Biol.* **249**, 624–632 (2007).

37. Birkhead, T. R., Briskie, J. V. & Møller, A. P. Male sperm reserves and copulation frequency in birds. *Behav. Ecol. Sociobiol.* **32**, 85–93 (1993).

38. Moller, A. P. Sperm competition, sperm depletion, paternal care, and relative testis size in birds. *Am. Nat.* **137**, 882–906 (1991).

39. Birkhead, T. R. & Montgomerie, R. Three decades of sperm competition in birds. *Phil. Trans. R. Soc. B* **375**, 20200208 (2020).

40. Brouwer, L. & Griffith, S. C. Extra-pair paternity in birds. *Mol. Ecol.* **28**, 4864–4882 (2019).

41. Hunter, F. M., Harcourt, R., Wright, M. & Davis, L. S. Strategic allocation of ejaculates by male Adélie penguins. *Proc. R. Soc. Lond. B* **267**, 1541–1545 (2000).

42. Hamamah, S. & Gatti, J. L. Role of the ionic environment and internal pH on sperm activity. *Hum. Reprod.* **13**, 20–30 (1998).

43. Gribbins, K. Reptilian spermatogenesis. *Spermatogenesis* **1**, 250–269 (2011).

44. Gribbins, K. M., Gist, D. H. & Congdon, J. D. Cytological evaluation of spermatogenesis and organization of the germinal epithelium in the male slider turtle, *Trachemys scripta*. *J. Morphol.* **255**, 337–346 (2003).

45. Schulz, R. W. et al. Spermatogenesis in fish. *Gen. Comp. Endocrinol.* **165**, 390–411 (2010).

46. Lubzens, E., Young, G., Bobe, J. & Cerdà, J. Oogenesis in teleosts: how fish eggs are formed. *Gen. Comp. Endocrinol.* **165**, 367–389 (2010).

47. Jalabert, B. Particularities of reproduction and oogenesis in teleost fish compared to mammals. *Reprod. Nutr. Dev.* **45**, 261–279 (2005).

48. Jónsson, H. et al. Multiple transmissions of de novo mutations in families. *Nat. Genet.* **50**, 1674–1680 (2018).

49. Martin, H. C. et al. Insights into platypus population structure and history from whole-genome sequencing. *Mol. Biol. Evol.* **35**, 1238–1252 (2018).

50. Smeds, L., Qvarnström, A. & Ellegren, H. Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **26**, 1211–1218 (2016).

51. Feng, C. et al. Moderate nucleotide diversity in the Atlantic herring is associated with a low mutation rate. *eLife* **6**, e23907 (2017).

52. Gao, Z., Wyman, M. J., Sella, G. & Przeworski, M. Interpreting the dependence of mutation rates on age and time. *PLoS Biol.* **14**, e1002355 (2016).

53. Goodman, M. Rates of molecular evolution: the hominoid slowdown. *BioEssays* **3**, 9–14 (1985).

54. Moorjani, P., Amorim, C. E. G., Arndt, P. F. & Przeworski, M. Variation in the molecular clock of primates. *Proc. Natl Acad. Sci. USA* **113**, 10607–10612 (2016).

55. Scally, A. & Durbin, R. Revising the human mutation rate: implications for understanding human evolution. *Nat. Rev. Genet.* **13**, 745–753 (2012).

56. Soojin, V. Y. Morris Goodman's hominoid rate slowdown: the importance of being neutral. *Mol. Phylogenet. Evol.* **66**, 569–574 (2013).

57. Faircloth, B. C. et al. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).

58. Garcia, J. A. & Lohmueller, K. E. Negative linkage disequilibrium between amino acid changing variants reveals interference among deleterious mutations in the human genome. *PLoS Genet.* **17**, e1009676 (2021).

59. Hedrick, P. W. & Garcia-Dorado, A. Understanding inbreeding depression, purging, and genetic rescue. *Trends Ecol. Evol.* **31**, 940–952 (2016).

60. Bonnet, T. et al. Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. *Science* **376**, 1012–1016 (2022).

61. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).

62. Chen, Y. et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* **7**, 1–6 (2017).

63. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

64. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at *bioRxiv* https://doi.org/10.1101/201178 (2018).

65. Kong, A. et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).

66. Besenbacher, S. et al. Novel variation and de novo mutation rates in population-wide de novo assembled Danish trios. *Nat. Commun.* **6**, 5969 (2015).

67. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).

68. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

69. Alfaro, M. E. et al. Explosive diversification of marine fishes at the Cretaceous–Palaeogene boundary. *Nat. Ecol. Evol.* **2**, 688–696 (2018).

70. Faircloth, B. C. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics* **32**, 786–788 (2016).

71. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

72. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

73. Sanderson, M. J. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* **19**, 101–109 (2002).

74. Kim, J. & Sanderson, M. J. Penalized likelihood phylogenetic inference: bridging the parsimony-likelihood gap. *Syst. Biol.* **57**, 665–674 (2008).

75. Meredith, R. W. et al. Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011).

76. Hughes, L. C. et al. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc. Natl Acad. Sci. USA* **115**, 6249–6254 (2018).

77. Benton, M. J. & Donoghue, P. C. J. Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* **24**, 26–53 (2007).

78. Green, R. E. et al. Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **346**, 1254449 (2014).

79. Sues, H. D. & Olsen, P. E. Triassic vertebrates of Gondwanan aspect from the Richmond basin of Virginia. *Science* **249**, 1020–1023 (1990).

80. Bauer, A. M., Böhme, W. & Weitschat, W. An Early Eocene gecko from Baltic amber and its implications for the evolution of gecko adhesion. *J. Zool.* **265**, 327–332 (2005).

81. Gelabert, P. et al. Evolutionary history, genomic adaptation to toxic diet, and extinction of the Carolina parakeet. *Curr. Biol.* **30**, 108–114.e5

(2020).

82. Maretty, L. et al. Sequencing and de novo assembly of 150 genomes from Denmark as a population reference. *Nature* **548**, 87–91 (2017).

83. Orme, D. et al. The caper package: comparative analysis of phylogenetics and evolution in R. R version 1.0.1 https://cran.r-project.org/package=caper (2018).

84. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).

85. Schmitz, J. et al. Genome sequence of the basal haplorrhine primate *Tarsius syrichta* reveals unusual insertions. *Nat. Commun.* **7**, 12997 (2016).

86. Vijay, N. et al. Population genomic analysis reveals contrasting demographic changes of two closely related dolphin species in the last glacial. *Mol. Biol. Evol.* **35**, 2026–2033 (2018).

87. Liu, Y. C. et al. Genome-wide evolutionary analysis of natural history and adaptation in the world's tigers. *Curr. Biol.* **28**, 3840–3849.e6 (2018).

88. Xu, S., Zhao, L., Xiao, S. & Gao, T. Whole genome resequencing data for three rockfish species of Sebastes. *Sci. Data* **6**, 97 (2019).

89. Yuan, Z. et al. Historical demography of common carp estimated from individuals collected from various parts of the world using the pairwise sequentially markovian coalescent approach. *Genetica* **146**, 235–241 (2018).

90. Fitak, R. R. & Johnsen, S. Green sea turtle (*Chelonia mydas*) population history indicates important demographic changes near the mid-Pleistocene transition. *Mar. Biol.* **165**, 110 (2018).

91. Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G. & Ellegren, H. Temporal dynamics of avian populations during pleistocene revealed by whole-genome sequences. *Curr. Biol.* **25**, 1375–1380 (2015).

92. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* **15**, 356 (2014).

93. Milholland, B. et al. Differences between germline and somatic mutation rates in humans and mice. *Nat. Commun.* **8**, 15183 (2017).

94. The 1000 Genomes Project. Variation in genome-wide mutation rates within and between human families. *Nat. Genet.* **43**, 712–714 (2011).

95. Rahbari, R. et al. Timing rates and spectra of human germline mutation. *Nat. Genet.* **48**, 126–133 (2015).

96. Wong, W. S. W. et al. New observations on maternal age effect on germline de novo mutations. *Nat. Commun.* **7**, 10486 (2016).

97. Turner, T. N. et al. Genomic patterns of de novo mutation in simplex autism. *Cell* **171**, 710–722.e12 (2017).

98. Sasani, T. A. et al. Large three-generation human families reveal post-zygotic mosaicism and variability in germline mutation accumulation. *eLife* **8**, e46922 (2019).

99. Kessler, M. D. et al. De novo mutations across 1465 diverse genomes reveal mutational insights and reductions in the Amish founder population. *Proc. Natl Acad. Sci. USA* **117**, 2560–2569 (2020).

100. Malinsky, M. et al. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat. Ecol. Evol.* **2**, 1940–1955 (2018).

101. Koch, E. M. et al. De novo mutation rate estimation in wolves of known pedigree. *Mol. Biol. Evol.* **36**, 2536–2547 (2019).

102. Harland, C. et al. Frequency of mosaicism points towards mutation-prone early cleavage cell divisions in cattle. Preprint at *bioRxiv* https://doi.org/10.1101/079863 (2017).

103. Pfeifer, S. P. Direct estimate of the spontaneous germ line mutation rate in African green monkeys. *Evolution* **71**, 2858–2870 (2017).

Acknowledgements

The authors thank the following contributors of samples for this study: A. Girard, C. Small, E. Couture, E. Gangloff, A. Bronikowski, F. Yu, H. Fernández, A. Carbajal Brossa, the Barcelona Zoo Biological Bank, J. Partecke, J. Judson, F. Janzen, J. Fjeldså, K. Thorup, K. Glover, L. Koren, M. Nagel, M. Fredholm, M. Liedvogel, T. Aquarium, P. Vullioud, S.-J. Luo, T. Gamble, Y. Yovel, J. Bakker, C. Bombis, T. Charlton, A. Corl, A. Foote, N. Geli, M. Guille, K. L. Hansen, W. Huizinga, M. Hunter, T. Knauf-Witzens, T. Lund Koch, S. Potier, A. Prahl, K. Robertson, C. Scala, M. Schellerup, I. Schnell, K. Vesterdorf, K. Wendelin, K. Worm and W.-z. Wang; G. Pacheco for valuable advice in the laboratory; K. Boomsma for stimulating conceptual discussions and for providing comments on the final version of this manuscript; and GenomeDK at Aarhus University for providing computational resources and support for this study. All sequencing data were generated with MGI-sequencers at the China National Genebank of BGI-Shenzhen. This project was funded by the Strategic Priority Research Programme (XDB13020000) and the International Partnership Programme (no. 152453KYSB20170002) of the Chinese Academy of Sciences, a Villum Investigator grant (no. 25900) from The Villum Foundation, and a Carlsberg Foundation Grant to G.Z. (CF16-0663). L.A.B. was supported by the Carlsberg Foundation and the Villum Foundation. M.H.S. was funded by the Novo Nordisk Foundation (NNF18OC0031004). J.I.H. was funded by the German Research Foundation (DFG) as part of the SFB TRR 212 (NC^3) (project nos. 316099922 and 396774617), the priority programme "Antarctic Research with Comparative Investigations in Arctic Ice Areas" SPP 1158 (project no. 424119118) and the sequencing costs in projects scheme (project no. 497640428).

Author information

Authors and Affiliations

1. Villum Centre for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Lucie A. Bergeron, Josefin Stiller & Guojie Zhang

2. Department of Molecular Medicine, Aarhus University, Aarhus, Denmark

Søren Besenbacher

3. BGI-Shenzhen, Shenzhen, China

Jiao Zheng & Panyi Li

4. BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, China

Jiao Zheng

5. Copenhagen Zoo, Frederiksberg, Denmark

Mads Frost Bertelsen

6. Parc Zoologique et Botanique de Mulhouse, Mulhouse, France

Benoit Quintard

7. Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany

Joseph I. Hoffman

8. British Antarctic Survey, High Cross, Cambridge, UK

Joseph I. Hoffman

9. College of Animal Science and Technology, Jilin Agricultural University, Changchun, China

Zhipeng Li

10. Department of Biomedical Sciences, Cornell University, Ithaca, NY, USA

Judy St. Leger

11. Key Lab of Sustainable Development of Marine Fisheries, Ministry of Agriculture and Rural Affairs, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, China

Changwei Shao

12. Center for Evolutionary Hologenomics, The GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

M. Thomas P. Gilbert

13. University Museum, NTNU, Trondheim, Norway

M. Thomas P. Gilbert

14. Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark

Mikkel H. Schierup

15. Centre for Evolutionary & Organismal Biology, Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China

Guojie Zhang

16. Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou, China

Guojie Zhang

17. State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

Guojie Zhang

Contributions

G.Z., M.H.S., S.B. and L.A.B. conceived this work. M.F.B., B.Q., J.I.H., Z.L., J.S.L. and C.S. provided samples for several species, as well as input into the writing and results interpretation. L.A.B., J.Z., P.L. and M.T.P.G. participated in the extraction, library preparation and sequencing. All analyses were conducted by L.A.B. with input from J.S. for the phylogenetic analysis and S.B. for the mutation rate estimation. L.A.B., G.Z., S.B., M.H.S. and J.I.H. wrote the initial draft of the manuscript with input from all co-authors. G.Z. supervised this project.

Corresponding authors

Correspondence to Lucie A. Bergeron or Guojie Zhang.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Anne Goriely and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Association of parental ages.

Maternal and paternal ages are significantly positively correlated for the 105 trios with known parental age at reproduction (linear regression; adjusted $r^2 = 0.77$, $F = 342.3$ on 1 and 103 DF, $p < 2.2 \times 10^{-16}$).

Extended Data Fig. 2 Comparison of published male bias estimates (α) using genome alignments and our male bias estimates (modified Fig. 1c of the main text).

The yellow points are α estimates from Wilson Sayres et al.[28], and the purple points are α estimates from Wu et al.[31]. Most of the common species reveal similar estimates with overlapping 95% confidence intervals. However, the estimates of α based on genome alignments are generally lower for dogs and cats than our estimates, yet the pedigree-based estimate of α for cats (Wang et al.[20]; green point) is similar to our estimate. See also Supplementary Table 5. The barplots represent male biases estimated by clustering different species per group (to have a minimum of 30 phased mutations per group) and the 95% confidence intervals were based on the binomial distribution. The silhouette of *Sygnathus* was created by J.S. All other silhouettes are from PhyloPic (http://phylopic.org), except one of the silhouettes of *Sarcophilus harrissi*, which was created by S. Werning, and the silhouette of *Pan troglodytes*, which was created by T. M. Keesey (vectorization) and T. Hisgett (photography); both are available under a CC-BY 3.0 licence (https://creativecommons.org/licenses/by/3.0).

Extended Data Fig. 3 Robustness of the calibration.

We compared the estimated substitution rates using the 14 initial calibration points with the inferred substitution rates using only 13 calibration points (with 14 iterations to remove each calibration node one by one). We found a strong relationship between the rates estimated with 14 and 13 calibrations (linear regression adjusted $r^2 = 0.91$, $F = 9416$ on 1 and 950 DF, p-value: $< 2.2 \times 10^{-16}$). However, some of the calibration points had a stronger impact on the estimated substitution rates. For instance, removing the two bird nodes (7 and 10), the gekko node (9), the Canidae/Arctoidea node (13) and the Glires/Primate node (8) altered some of the substitution rate estimates.

Extended Data Fig. 4 Per-generation mutation rates (similar to Fig. 1a) including published data on closely related species.

For each species, the colored squares represent the average per-generation observed rate, along with the 95% confidence intervals based on the binomial distribution, and the black points represent published estimates from similar or closely related species to those included in our dataset. For most of the species, these estimates lie within the 95% confidence intervals of our estimates. Published estimates are from: *Felis catus* (Wang et. al.[20]), *Mus musculus* (Milholland et al.[93], Lindsay et al.[34]), *Pan troglodytes* (Venn et al.[21], Tatsumoto et al.[23], Besenbacher et al.[16]), *Homo sapiens* (Conrad et al.[94], Kong et al.[65], Francioli et al.[32], Rahbari et al.[95], Wong et al.[96], Jónsson et al.[22], Maretty et al.[82], Turner et al.[97], Sasani et al.[98], Kessler et al.[99]). The closely related species are from: close to the *Salmo salar*, *Clupea harengus* (Feng et al.[51]), close to *Paralichthys olivaceus*, the Cichlid (Malinsky et al.[100]), close to *Canis lupus familiaris*, *Canis lupus* (Koch et al.[101]), close to *Capra hircus*, *Bos taurus* (Harland et al.[102]), close to *Mandrillus leucophaeus*, *Papio anubis* (Wu et al.[13]), *Macaca mulatta* (Wang et al.[14], Bergeron et al.[12]), and *Chlorocebus sabaeus* (Pfeifer[103]), close to *Saimiri boliviensis boliviensis*, *Aotus nancymaae* (Thomas et al.[17]), close to *Monodelphis domestica*, *Ornithorhynchus anatinus* (Martin et al.[49]), close to *Taeniopygia guttata*, *Ficedula albicollis* (Smeds et al.[50]). See also Supplementary Table 8. The silhouette of *Sygnathus* was created by J.S. All other silhouettes are from PhyloPic (http://phylopic.org), except one of the silhouettes of *S. harrissi*, which was created by S. Werning, and the silhouette of *P. troglodytes*, which was created by T. M. Keesey (vectorization) and T. Hisgett (photography); both are available under a CC-BY 3.0 licence (https://creativecommons.org/licenses/by/3.0).

Extended Data Fig. 5 Germline mutation rates are associated with long-term substitution rates.

This figure is similar to the main Fig. 2 but uses phylogenetic regression (PGLS) on a log scale. The grey dashed lines indicate equality. **a**. Using a log scale, there is a significant positive correlation between the per-year rates and the rates derived from Ultraconserved elements (UCEs) and their

flanking sequences. **b**. However, this correlation is not significant when comparing the per-year rates with the rates derived from the whole genome alignments (WGAs).

Extended Data Fig. 6 Comparison of substitution rates estimated with Ultra Conserved Elements (UCEs) and MultiZ alignments.

The substitution rates estimated with the two methods are highly correlated (linear regression: adjusted $r^2 = 0.73$, F = 179.9 on 1 and 66 DF, $p < 2.2 \times 10^{-16}$).

Extended Data Fig. 7 Three life-history traits are not significantly associated with the per-generation mutation rate.

a. lifespan in the wild, **b**. body mass and **c**. the mating system (polygamy versus monogamy). The total number of species with modeled per-generation rate was 55 for the phylogenetic regression (PGLS). The boxplots represent the median, the interquartile range, and the maximum and minimum excluding outliers.

Extended Data Fig. 8 The drift barrier hypothesis on different times and different mutation rate parameters used to estimate N_e with phylogenetic regression (PGLS).

a. The correlation between N_e and the mutation rate per generation is not significant when using the most recent value before 30,000 years estimated by PSMC. **b**. The relationship is also not significant when using the harmonic mean over a more recent period of time (30,000 years to 130,000 years ago). However, this relationship is significant for mammals (adjusted $r^2 = 0.104$, $p = 0.04$). We used the harmonic mean over the past million years in the main text, as PSMC is not reliable over recent periods. **c**. When looking at the relationship between the mutation rate and N_e, estimated using the pedigree-based mutation rate, we find a stronger signal over the past 1,000,000 years, probably due to the circularity of this analysis. **d**. However,

the relationship is still not significant when using the most recent time point or **e**. the average over the past 100,000 years.

Extended Data Fig. 9 Effective population sizes calculated with two different methods (see main text) are significantly correlated.

The harmonic mean of the population size estimated with PSMC from 30,000 to 1,000,000 years ago is significantly correlated with the effective population size estimated from $N_e = \pi/4\mu$ (linear regression: adjusted $r^2 = 0.83$, $F = 316.3$ on 1 and 63 DF, $p < 2.2 \times 10^{-16}$).

Supplementary information

Supplementary Figures

This file contains the Supplementary Figs. 1–8.

Reporting Summary

Peer Review File

Supplementary Note

This file contains detailed information on the model used to estimate germline mutation rate per generation at generation time.

Supplementary Tables

This file contains Supplementary Tables 1–9.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation,

distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Bergeron, L.A., Besenbacher, S., Zheng, J. *et al.* Evolution of the germline mutation rate across vertebrates. *Nature* **615**, 285–291 (2023). https://doi.org/10.1038/s41586-023-05752-y

- Received: 19 November 2021
- Accepted: 23 January 2023
- Published: 01 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05752-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

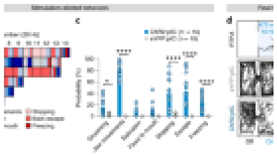
Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Rate at which mutations are passed to offspring established for various animal species

Nature Research Briefing 01 Mar 2023

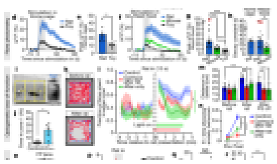
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05752-y



Aversive state processing in the posterior insular cortex

27 August 2019

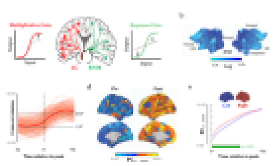
Daniel A. Gehrlach, Nate Dolensek, … Nadine Gogolla



Stimulus-specific hypothalamic encoding of a persistent defensive state

16 September 2020

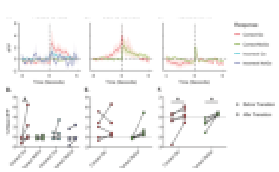
Ann Kennedy, Prabhat S. Kunwar, … David J. Anderson



The ascending arousal system shapes neural dynamics to mediate awareness of cognitive states

14 October 2021

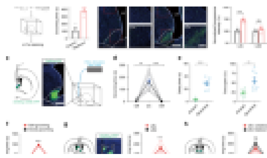
Brandon R. Munn, Eli J. Müller, … James M. Shine



[Optogenetic and pharmacological interventions link hypocretin neurons to impulsivity in mice]

19 January 2023

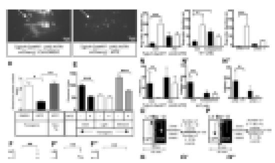
Susan M. Tyree, Kimberly J. Jennings, … Luis de Lecea



[A limbic circuitry involved in emotional stress-induced grooming]

08 May 2020

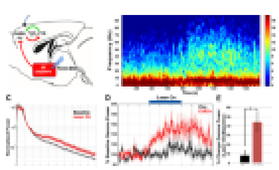
Ming-Dao Mu, Hong-Yan Geng, … Ya Ke



[Brain-wide perception of the emotional valence of light is regulated by distinct hypothalamic neurons]

28 April 2022

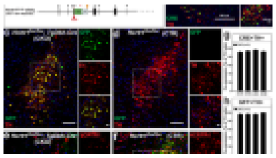
Mahendra Wagle, Mahdi Zarei, … Su Guo



[Optogenetic manipulation of an ascending arousal system tunes cortical broadband gamma power and reveals functional deficits relevant to schizophrenia]

20 July 2020

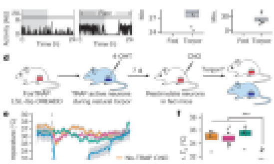
James M. McNally, David D. Aguilar, … Ritchie E. Brown



Bidirectional and context-dependent changes in theta and gamma oscillatory brain activity in noradrenergic cell-specific Hypocretin/Orexin receptor 1-KO mice

19 October 2018

Sha Li, Paul Franken & Anne Vassalli



Neurons that regulate mouse torpor

11 June 2020

Sinisa Hrvatin, Senmiao Sun, … Michael E. Greenberg

- Article
- Open Access
- Published: 01 March 2023

Cardiogenic control of affective behavioural state

- Brian Hsueh,
- Ritchie Chen,

- YoungJu Jo,
- Daniel Tang,
- Misha Raffiee,
- Yoon Seok Kim,
- Masatoshi Inoue,
- Sawyer Randles,
- Charu Ramakrishnan,
- Sneha Patel,
- Doo Kyung Kim,
- Tony X. Liu,
- Soo Hyun Kim,
- Longzhi Tan,
- Leili Mortazavi,
- Arjay Cordero,
- Jenny Shi,
- Mingming Zhao,
- Theodore T. Ho,
- Ailey Crow,
- Ai-Chi Wang Yoo,
- Cephra Raja,
- Kathryn Evans,
- Daniel Bernstein,
- Michael Zeineh,
- Maged Goubran &
- …
- Karl Deisseroth

Show authors

Nature **volume 615**, pages 292–299 (2023)

- 32k Accesses
- 763 Altmetric
- Metrics details

Abstract

Emotional states influence bodily physiology, as exemplified in the top-down process by which anxiety causes faster beating of the heart[1,2,3]. However, whether an increased heart rate might itself induce anxiety or fear responses is unclear[3,4,5,6,7,8]. Physiological theories of emotion, proposed over a century ago, have considered that in general, there could be an important and even dominant flow of information from the body to the brain[9]. Here, to formally test this idea, we developed a noninvasive optogenetic pacemaker for precise, cell-type-specific control of cardiac rhythms of up to 900 beats per minute in freely moving mice, enabled by a wearable micro-LED harness and the systemic viral delivery of a potent pump-like channelrhodopsin. We found that optically evoked tachycardia potently enhanced anxiety-like behaviour, but crucially only in risky contexts, indicating that both central (brain) and peripheral (body) processes may be involved in the development of emotional states. To identify potential mechanisms, we used whole-brain activity screening and electrophysiology to find brain regions that were activated by imposed cardiac rhythms. We identified the posterior insular cortex as a potential mediator of bottom-up cardiac interoceptive processing, and found that optogenetic inhibition of this brain region attenuated the anxiety-like behaviour that was induced by optical cardiac pacing. Together, these findings reveal that cells of both the body and the brain must be considered together to understand the origins of emotional or affective states. More broadly, our results define a generalizable approach for noninvasive, temporally precise functional investigations of joint organism-wide interactions among targeted cells during behaviour.

Main

Interoceptive processing of visceral physiological signals, such as cardiac palpitations or stomach fullness, is crucial for maintaining homeostasis[1,2,3]. Diverse psychiatric conditions, such as anxiety disorders, panic disorder, body dysmorphic disorders and addiction, have been hypothesized to be related to dysregulation of interoceptive monitoring by the brain[3,4], and can be statistically correlated with specific visceral organ dysfunction. For example, patients with panic disorder and agoraphobia are more likely to have mitral valve prolapse or clinical symptoms similar to paroxysmal

supraventricular tachycardia[5,6]. Modern correlative studies have further suggested links between cardiac changes and affect regulation[7,8], including correlations between cardiac interoception with anxiety and functional alterations in the insular cortex, a cortical region that has a central role in both the processing of physiological signals and the regulation of emotions[4,10]. However, determining whether primary physiological signals such as increased heart rate can causally influence behavioural states, as proposed in classical physiological theories of emotion[9], has—although widely debated—remained largely experimentally intractable[11]. Available nonspecific interventions that might disrupt cardiac signals (such as electrical vagus nerve stimulation) are well known to also induce numerous physiological changes that would be unwanted in this context, including direct suppression of respiratory and heart rates as well as anxiolytic and antidepressive effects[11,12,13], giving rise to multiple direct confounds for the question explored here. Other nonspecific interventions to alter cardiac rhythms, such as broadly active pharmacological stimulants or electrical pacemakers[14], also introduce insuperable confounds through initial actions beyond the direct pacing of cardiomyocytes, and thus lack the necessary precision. Studying the key question of how cardiac physiology regulates emotional states has remained inaccessible, and the effects on behaviour remain unknown.

Precise modulation of electrochemical signals in the heart and other peripheral organs in vivo would enable fundamental studies of physiology and interoceptive signalling[15,16,17,18,19], but stimulation methods that operate with high spatial and temporal precision in highly dynamic environments such as the beating heart[20,21,22,23,24,25,26,27,28] are limited. Electrical pacemakers require invasive surgical implantation to deliver local indiscriminate stimulation that lacks cell-type specificity[20,21,22,23]. Optogenetics might in principle facilitate cardiomyocyte-specific control with high spatial and temporal precision[14], but existing optogenetic methods have been limited to acute demonstrations that require exposure or even excision of the heart to deliver light[23,24,25,26,27,28], all of which are incompatible with freely moving studies of behaviour. Thus far, to our knowledge, no study beyond the brain has achieved precise and noninvasive organ-level control of behavioural or physiological function. Establishing noninvasive approaches to manipulate physiology with cell-type specificity

would enable long-sought functional studies of signals that arise from cells across the organism, and reveal the causal influences of these cells on brain function and behaviour.

Conventional microbial opsins have not been sensitive enough to control a large organ such as the heart with the requisite power to facilitate behavioural studies within intact animals[23,24,25,26,27,28], but the discovery of the highly sensitive and red-shifted pump-like channelrhodopsin ChRmine led us to consider the potential for noninvasive optogenetic control of deep tissue with minimal irradiance[29]. We previously found that ChRmine enabled neuromodulation of deep brain circuits without intracranial surgery[30], which raised the possibility that this optogenetic tool might be broadly applicable to modulating biological processes across the entire body of large organisms such as mammals. Specifically, we hypothesized that ChRmine might allow on-demand deep-tissue control of cardiac pacing when targeted to cardiomyocytes without the need for direct exposure of the heart[29,30].

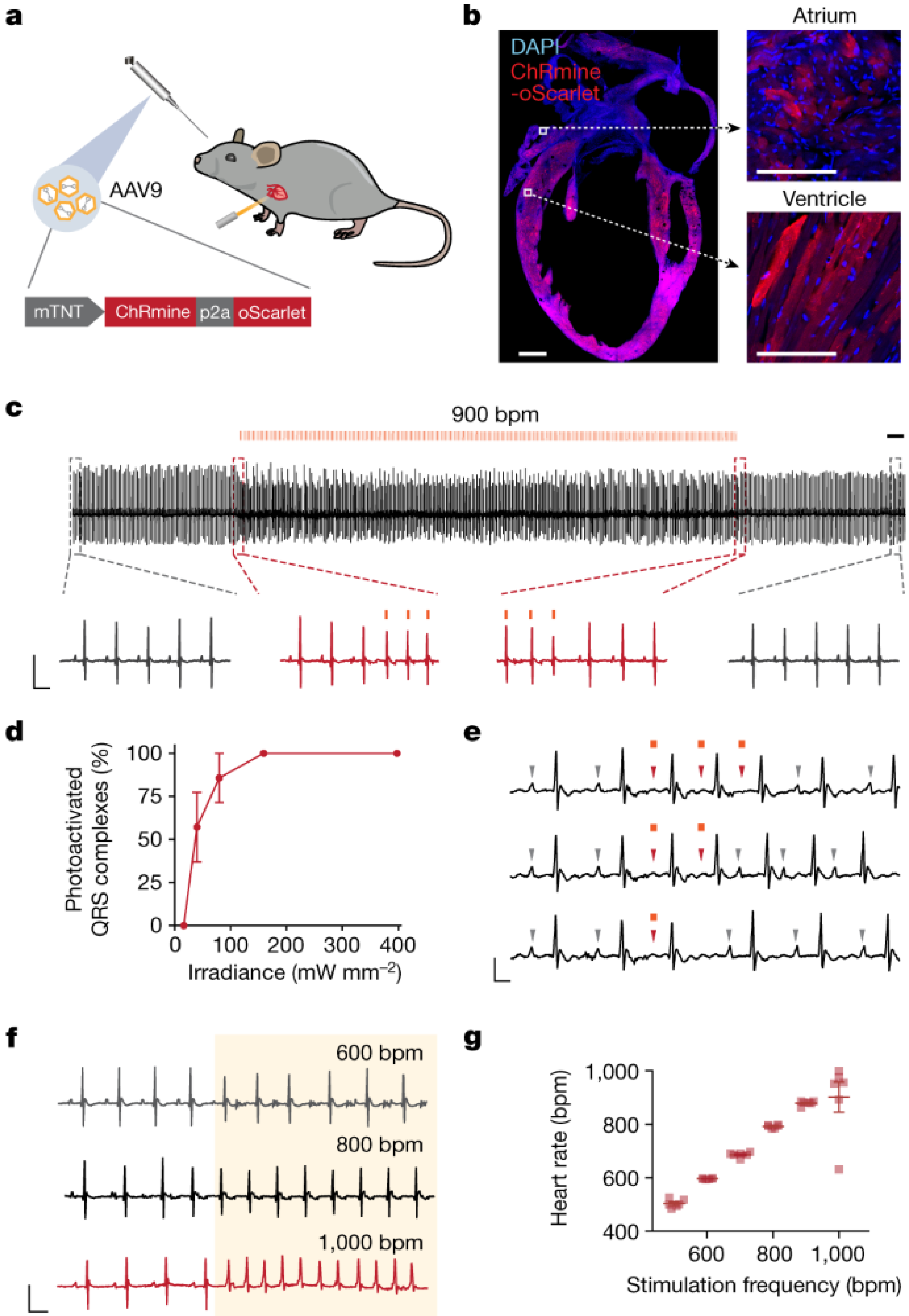
Noninvasive optogenetic cardiac control

We first achieved cardiomyocyte-restricted expression by placing the ChRmine transgene under the control of the mouse cardiac troponin T promoter (mTNT), using the AAV9 serotype, which exhibits tropism for cardiac tissue[25,31]. Infection of cultured primary cardiomyocytes with AAV9-mTNT::ChRmine-2A-oScarlet enabled light-evoked contractions with irradiance as low as 0.1 mW mm^{-2}, consistent with the photosensitivity of ChRmine in neurons[29] (Extended Data Fig. 1 and Supplementary Video 1).

We next determined whether systemic viral gene delivery of ChRmine, despite the lower multiplicity of infection compared with transduction by direct local injection[30,32], could allow noninvasive control of heart rhythms in wild-type mice. Retro-orbital injection of AAV9 enabled restricted expression of ChRmine in cardiomyocytes throughout the heart, with homogeneous expression in both ventricular and atrial walls and no off-target expression in other cardiac cell types (fibroblasts and neuronal

ganglia) or in other organs (Fig. 1a,b and Extended Data Fig. 2). When pulsed 589-nm light was delivered through intact skin overlying the thorax of anaesthetized mice, we observed robust photoactivation of cardiac QRS complexes within a safe range of irradiance comparable to that used for transcranial optogenetics[30] (Fig. 1c,d). Reliable cardiac rhythms were induced at even supraphysiological rates of up to 900 beats per minute (bpm), within the photophysical properties of ChRmine[33], with an immediate return to naturally paced sinus rhythm upon the cessation of light (Fig. 1e–g). This approach was not able to decrease heart rate below baseline levels, but afforded spatial control of cardiac rhythms by evoking either right or left ventricular pacing depending on the placement of the laser (Extended Data Fig. 3a–i).

Fig. 1: Development of a noninvasive optical pacemaker.

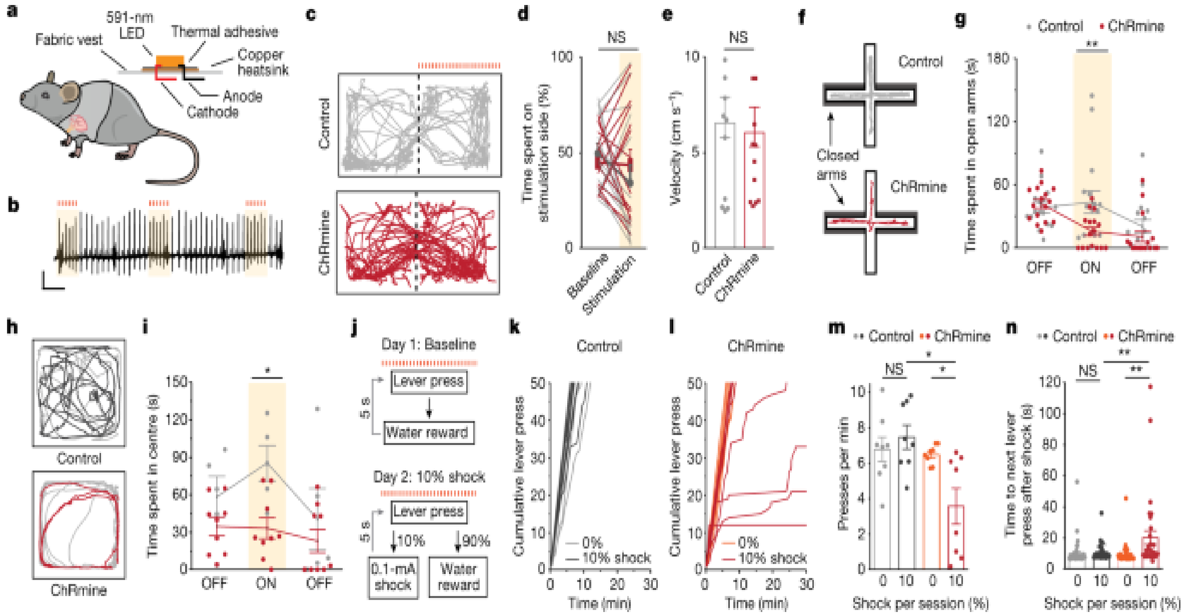


a
AAV9
mTNT
ChRmine
p2a
oScarlet
b
DAPI
ChRmine
-oScarlet
Atrium
Ventricle
c
900 bpm
d
Photoactivated QRS complexes (%)
100
75
50
25
0
0
100
200
300
400
Irradiance (mW mm^{-2})
e
f
600 bpm
800 bpm
1,000 bpm
g
Heart rate (bpm)
1,000
800
600
400
600
800
1,000
Stimulation frequency (bpm)

a, Schematic showing the optical control of cardiac rhythm with an external light source enabled by retro-orbital injection of AAV9-mTNT::ChRmine-p2A-oScarlet. **b**, Confocal cross-section images indicating homogeneous transgene expression of ChRmine-p2A-oScarlet (red) with DAPI staining (blue) in atria and ventricles. Scale bars, 1 mm (main); 100 µm (inset). **c**, Example electrocardiogram (ECG) trace with optical pacing using 589 nm light delivered at 15 Hz (at 900 bpm) with a pulse width of 10 ms and irradiance of 160 mW mm^{-2}. Scale bar, 500 ms. Inset traces: ECG signal before and after light delivery (grey) and at light onset and cessation (red). Scale bar, 50 ms, 0.5 mV. **d**, Reliability of photoactivated QRS complexes at 900 bpm as a function of cutaneous optical irradiance ($n = 6$ mice). **e**, Example ECG traces of individual 10-ms optical pulses. Grey arrowheads indicate P waves associated with sinus rhythm, which are overridden (red arrowheads) during optical pacing. Scale bar, 25 ms, 0.25 mV. **f**, Example ECG traces of pacing at 600, 800 and 1,000 bpm. Scale bar, 50 ms, 0.5 mV. The shaded area indicates the period of illumination at the specified frequency at 100% duty cycle. **g**, Characterization of optical pacing fidelity, showing stimulation frequency versus ECG-measured heart rate ($n = 6$ mice). All ECG measurements were performed in anaesthetized mice. Data are mean ± s.e.m.

To translate this approach to mouse behaviour, we mounted a 591-nm micro-LED onto a wearable fabric vest to deliver light through intact skin overlying the chest wall (Fig. 2a and Extended Data Fig. 4a,b). This integration of a molecular tool with accessible electronics enabled the initial demonstration of noninvasive and sustained ventricular pacing at experimenter-defined rhythms suitable for most behavioural assays in freely moving mice (Extended Data Fig. 4).

Fig. 2: Optically induced tachycardia increases anxiety-like behaviour.



a, Schematic of a micro-LED mounted to a wearable vest and fastened onto a mouse. **b**, Representative ECG trace of optically induced tachycardia (900 bpm for 500 ms every 2 s) used for all behavioural assays. Scale bar, 0.2 mV, 500 ms. **c**, Example path trace of a mouse with (red) or without (grey) ChRmine expression during an RTPP test, in which mice received optically induced cardiac pacing on one side of the chamber. **d**, Percentage of time spent on stimulation side during baseline and stimulation days for control (grey) and ChRmine-expressing (red) mice ($n = 16$ mice per group; two-way repeated-measures ANOVA with Bonferroni post hoc test: group (opsin) × time interaction $F_{(1,30)} = 2.29$, $P = 0.14$; group (opsin) effect $F_{(1,30)} = 6.2 \times 10^{-4}$, $P = 0.98$; time effect $F_{(1,30)} = 2.06$, $P = 0.16$. Bonferroni post hoc: control versus ChRmine $P = 0.71$ (baseline day); $P = 0.67$ (stimulation day)). NS, not significant. **e**, Average velocity on the optically paced side during RTPP ($n = 16$ mice per group; unpaired two-tailed *t*-test, $P = 0.81$). **f**, Example path trace of control (grey) and ChRmine-expressing (red) mice during an EPM test with optical pacing during the 5-min ON epoch of a 15-min trial. Open arms are vertical; closed arms are horizontal and bordered in grey. **g**, Time spent in open arms during 5-min epochs of EPM exploration ($n = 16$ mice per group; two-way repeated-measures ANOVA with Bonferroni post hoc test: group (opsin) × time interaction $F_{(2,60)} = 3.906$, $P = 0.0254$; group (opsin) effect $F_{(1,30)} = 3.297$, $P = 0.0794$; time effect $F_{(2,60)} = 9.75$, $P = 0.0002$. Bonferroni post hoc: ON epoch ChRmine versus control,

$P = 0.0079$). **h, Example path trace of control (grey) and ChRmine-expressing (red) mice during an OFT with optical pacing during the 3-min ON epoch of a 9-min trial. **i**, Time spent in the centre during 3-min epochs of OFT exploration ($n = 5$ (control), 9 (ChRmine) mice; two-way repeated-measures ANOVA with Bonferroni post hoc test: group (opsin) × time interaction $F_{(2,24)} = 1.531$, $P = 0.024$; group (opsin) effect $F_{(1,12)} = 5.69$, $P = 0.0035$; time effect $F_{(2,24)} = 3.42$, $P = 0.049$. Bonferroni post hoc: ON epoch ChRmine versus control, *$P = 0.018$). **j**, Vogel conflict task to assay for cardiogenic effects on operant behaviour. Water-restricted mice were first trained for 2–3 weeks until each mouse was able to complete the 50 water-reward lever-press trials over 30 min for at least 3 consecutive days. On the day of the behavioural task, mice received optical pacing while completing a total of 50 lever-press trials per session (day 1). On the subsequent day, a 10% pseudorandom chance of shock was introduced upon lever press (day 2). **k**,**l**, Cumulative lever presses during 0% (day 1) and 10% shock (day 2) sessions for control (**k**) and ChRmine-expressing (**l**) mice ($n = 8$ mice). **m**, Average lever-pressing rate during 0% baseline or 10% shock trial sessions ($n = 8$ mice per group; two-way repeated-measures ANOVA with Bonferroni post hoc test: group (opsin) × condition (shock) interaction $F_{(1,14)} = 8.326$, $P = 0.0120$; group (opsin) effect $F_{(1,14)} = 7.39$, $P = 0.0166$; condition (shock) effect $F_{(1,14)} = 3.162$, $P = 0.0971$. Bonferroni post hoc: control 0% versus 10%, $P = 0.8933$; ChRmine 0% versus 10%, *$P = 0.0106$; 0% control versus ChRmine, $P > 0.9999$; 10% control versus ChRmine, **$P = 0.0010$). **n**, Elapsed time between a lever press resulting in shock and a subsequent lever press as a measure of the mouse's apprehension state. On 10% shock trials, 3 out of 8 mice did not complete the trial, so the time to next lever press for some trials cannot be measured ($n = 40$, 40, 40 and 32 presses per group in control 0%, control 10%, ChRmine 0% and ChRmine 10%; two-way ANOVA with Bonferroni post hoc test: group (opsin) × condition (shock) interaction $F_{(1,148)} = 6.478$, $P = 0.0119$; group (opsin) effect $F_{(1,148)} = 5.041$, $P = 0.0262$; condition (shock) effect $F_{(1,148)} = 7.253$, $P = 0.0079$. Bonferroni post hoc: control 0% versus 10%, $P > 0.9999$; ChRmine 0% versus 10%, **$P = 0.0026$; 0% control versus ChRmine, $P > 0.9999$; 10% control versus ChRmine, **$P = 0.0074$). Data are mean ± s.e.m.

To test whether heart rhythms directly set by this optical pacemaker could influence behaviour, we optogenetically induced intermittent ventricular tachycardia (900 bpm for 500 ms every 1,500 ms) to mimic non-sustained arrhythmias that are observed during stressful contexts[34,35,36], while shortening the duration of decreases in systolic blood pressure and avoiding incidental heating from light-delivery devices (Fig. 2b and Extended Data Figs. 3j–m and 4c). We first assessed the appetitive or aversive effects of optical pacing using a real-time place-preference (RTPP) assay (Fig. 2c). Mice spent an equal proportion of time on the paced and non-paced sides of the two-chamber arena, and showed no difference in locomotion compared to littermate controls—revealing that optically induced intermittent tachycardia was not intrinsically aversive and did not cause locomotor impairment (Fig. 2d,e). Optical pacing also did not modulate pain perception during a hot-plate test, with paced mice exhibiting comparable behavioural responses to control mice (Extended Data Fig. 5a–c).

Anxiety-like state evoked by cardiac pacing

By contrast, when we tested for anxiety-related behaviour using an elevated plus maze (EPM) assay, the same mice exhibited limited exploration of the open (exposed) arms of the apparatus after optical pacing, compared to control mice, preferring to remain within the protected areas of the closed arms (Fig. 2f,g). Paced mice also avoided the centre area during an open field test (OFT) (Fig. 2h,i). We observed no effects from illumination alone in control (opsin-negative) mice, and baseline anxiety levels between control and virally transduced groups were similar (Extended Data Fig. 5d–h). Increased anxiety-like behaviour induced by optical pacing during the EPM and OFT assays was similarly observed in female cohorts (Extended Data Fig. 5i–l). We found that mice that received intermittent cardiac pacing within baseline ranges (660 bpm) rather than elevated (900 bpm) ranges did not exhibit behavioural differences compared to control mice during the EPM or OFT (Extended Data Fig. 5m–p). Because continuous ventricular pacing can have a long-lasting effect on animal health[20,21], we also assessed for potential changes in baseline anxiety levels and mobility in mice that were subjected to longer-term treatments of intermittent tachycardia (one-hour sessions every other day for two weeks) and did not observe locomotor

or behavioural differences in these mice when compared to control mice during the EPM and OFT (Extended Data Fig. 5q–u).

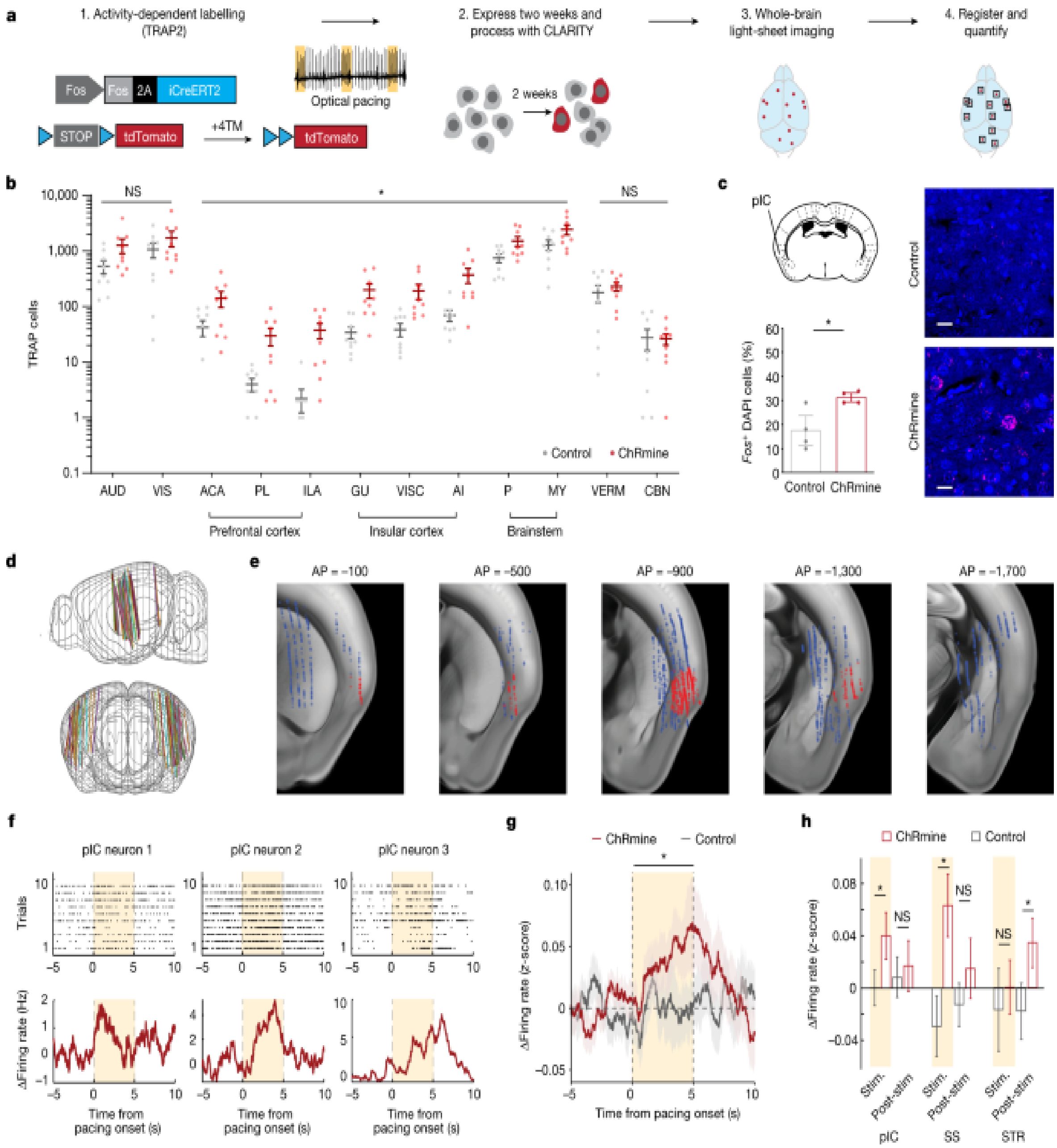
We further investigated whether this context-dependent enhancement of anxiety-related behaviour could translate to a classical operant task, by using a trial-based variation of the Vogel conflict task in which water-restricted mice show willingness to seek a water reward even when the reward is coupled to a risk of mild shock[37] (Fig. 2j). Mice that received cardiac pacing performed similarly to control littermates when allowed to freely press for water with no delivery of the aversive stimulus (Fig. 2k–n). However, when random shocks were introduced in 10% of trials, optically paced mice were found to suppress or terminate water-seeking altogether (Fig. 2l,m). These mice exhibited increased apprehension, as revealed both by an overall decreased lever-pressing rate and by an increased time to the next subsequent lever press after a shock trial—consistent with the heightened levels of anxiety that were observed during the EPM and OFT (Fig. 2m,n). By contrast, control mice exhibited reduced water-seeking only when the frequency of shock trials was increased to 30% (Extended Data Fig. 6). This context-dependent influence of cardiac pacing on anxiety-like behaviour suggested that higher-order brain function was involved in the processing of interoceptive cues.

Optical pacing increases insula activity

We therefore next used the optical pacemaker to identify potential neural correlates and mechanisms of this observed behaviour along the heart–brain axis. First, transgenic TRAP2 mice, in which neurons with increased expression of the immediate early gene *Fos* can be labelled with tdTomato as a marker for neural activation[38], were used to perform a brain-wide screen to identify regions that were affected by optical pacing (Fig. 3a,b and Extended Data Fig. 7). Whole-brain tissue clearing[39], registration to a reference brain atlas[40] and automated cell counting together revealed that a number of brain regions exhibited increased expression of tdTomato in optically paced mice. These included areas associated with the central autonomic network[41,42], such as the insular cortex (including its visceral area (VISC), gustatory area (GU) and agranular insular area (AI)), prefrontal cortex (including the infralimbic area (ILA), prelimbic area (PL) and

anterior cingulate area (ACA)) and brainstem (including the pons (P) and medulla (MY)) (Fig. 3b). Meanwhile, pointing to specificity, many other cortical regions that are not known to be involved in autonomic or interoceptive processing were not significantly activated, including primary sensory auditory (AUD) and visual (VIS) cortical areas, as well as the cerebellar vermis (VERM) and cerebellar nuclei (CBN) (Fig. 3b). Consistent with the TRAP2 mapping results, optical pacing similarly increased the endogenous expression of *Fos* mRNA in the posterior insular cortex (pIC) and in the brainstem (Fig. 3c and Extended Data Fig. 8). In particular, sensory relay circuits of the nucleus tractus solitarius (NTS), as well as noradrenergic neurons in the locus coeruleus (LC), which are involved in arousal and stress[43], also exhibited prominent *Fos* labelling (Extended Data Fig. 8).

Fig. 3: Whole-brain screen for regions that are activated by optically induced tachycardia.



a, Schematic for whole-brain activity mapping to identify regions that are activated by optically paced tachycardia. Double-transgenic TRAP2;Ai14 reporter mice were injected with 4-hydroxytamoxifen (4TM) and treated with optically induced tachycardia for 15 min. After two weeks of tdTomato reporter gene expression, mice were euthanized and processed with CLARITY. Whole brains were imaged with a light-sheet microscope, followed by automated registration to a common brain atlas, cell segmentation and quantification to identify brain regions with differential accumulation of activated TRAP (tdTomato^{+}) cells. **b**, Regional cell counts

of paced (ChRmine, red) versus control (grey) cohorts sorted from anterior to posterior anatomical regions with select regions from the central autonomic network with increased TRAP cells and regions outside of the central autonomic network without statistical significance (*n* = 9 per group; multiple two-sided *t*-tests corrected for multiple comparisons with the Benjamini and Hochberg method (*false discovery rate (FDR) = 10%)). Significantly activated regions include the prefrontal cortex (ACA, PL and ILA), insular cortex (GUI, VISC and AI) and brainstem (P and MY). Non-significant regions include primary sensory cortices (AUD, VIS) and the cerebellum (VERM, CBN). **c**, Percentage of cells that are Fos^+, determined from in situ hybridization for *Fos* mRNA (magenta) after cardiac pacing in control and ChRmine-expressing mice in the pIC (*n* = 4 mice per group; unpaired two-tailed *t*-test, *$P = 0.020$). Scale bars, 20 µm. **d**, Electrode tracks from *n* = 5 mice (3 ChRmine and 2 control) over 60 recording sessions co-registered to the common Allen Brain Atlas. **e**, Locations of recorded single units overlaid onto the Allen Brain Atlas. Red denotes units in the insular cortex. AP, anterior–posterior. **f**, Spike raster and changes in firing rate for three example insular neurons after 900-bpm pacing. **g**, Population-averaged changes in firing rate of insular neurons from ChRmine (red, *n* = 391) or control (grey, *n* = 228) mice (one-sided *P* values from hierarchical bootstrap: $P = 0.026$ (during 5 s pacing); $P = 0.357$ (during 5 s after pacing)). **h**, Average change in baseline firing rate per brain region across 5-s epochs during and after photostimulation in control (grey) and ChRmine-expressing (red) mice. Single units were obtained from the pIC (*n* = 228 (control), *n* = 391 (ChRmine)); somatosensory cortex (SS; *n* = 77 (control), *n* = 368 (ChRmine)); and striatum (STR; *n* = 70 (control), *n* = 800 (ChRmine)). One-sided *P* values from hierarchical bootstrap: pIC: $P = 0.026$ (stimulation; stim.), $P = 0.36$ (post-stimulation; post-stim.); SS: $P = 0.0014$ (stim.), $P = 0.16$ (post-stim.); STR: $P = 0.29$ (stim.), $P = 0.036$ (post-stim.). Data are mean ± s.e.m.

We next investigated the cardiac-pacing-induced neural dynamics at single-neuron resolution using in vivo electrophysiology in awake mice (Fig. 3d–h). On the basis of the observed increased *Fos* expression in the pIC after cardiac pacing, and because the insular cortex is a key cortical hub for interoception[44], we used four-shank Neuropixels 2.0 probes to obtain rich multi-regional recordings of the pIC and surrounding regions (Fig. 3d,e).

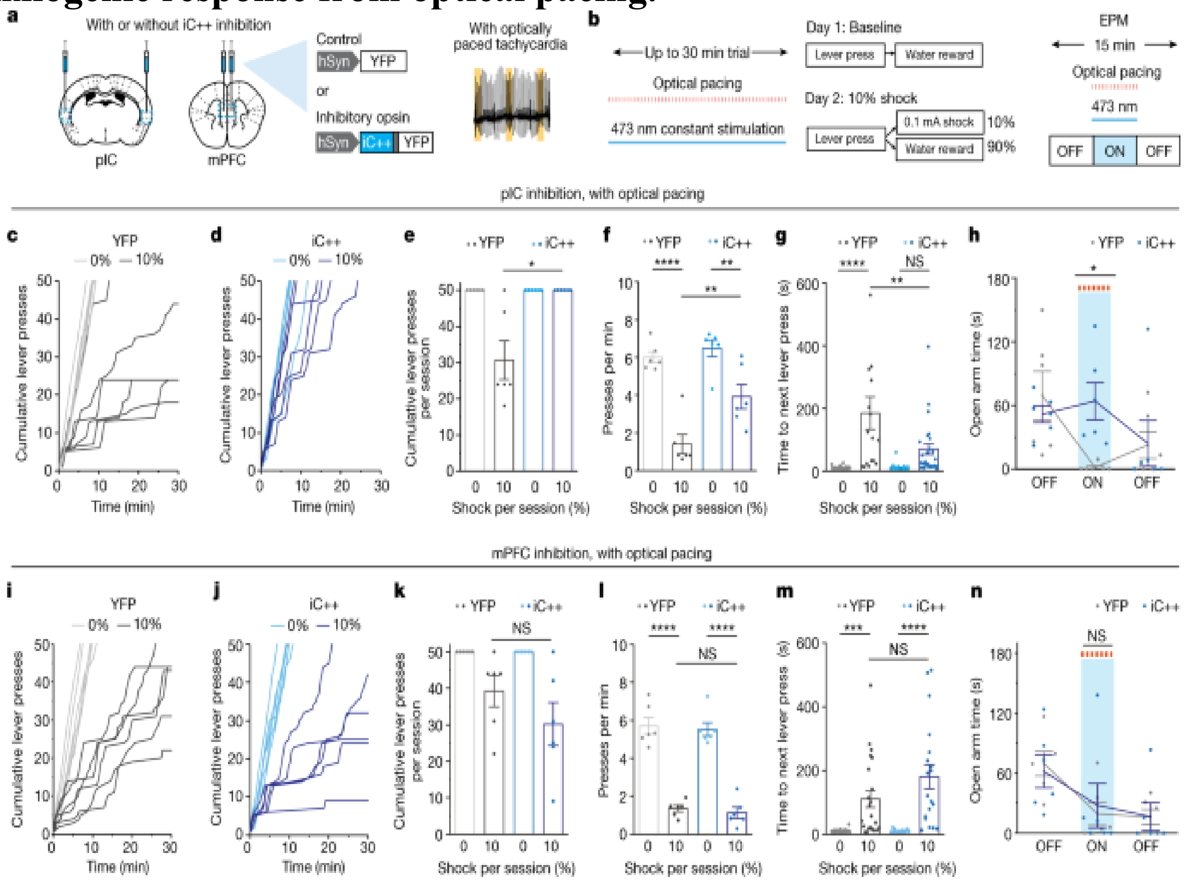
Notably, we observed pacing-evoked increases in insular activity at both the single-unit and population levels (Fig. 3f,g), whereby pIC neurons were acutely activated by cardiac stimulation with heterogeneous temporal dynamics and returned to basal levels of activity after pacing offset. By contrast, no significant changes in activity were recorded in control mice during photostimulation, ruling out potential light- or heat-induced representations in this region (Fig. 3g). We also observed a substantial diversity in temporal dynamics triggered by pacing in other recorded regions, including acute responses in the somatosensory cortex and delayed responses in the striatum during stimulation offset (Fig. 3h). Our data show that the pIC and other regions of the central autonomic network are distinctly engaged by optically evoked tachycardia[41,42], and are in line with human neuroimaging studies that have correlated these brain areas with cardiac interoception[45,46,47].

Insular inhibition reduces cardiac anxiety effects

To determine whether the anxiogenic circuitry recruited by cardiac pacing could be specifically modulated to influence behaviour, we next performed optogenetic inhibition with the 473 nm (blue light)-activated inhibitory channelrhodopsin iC++. We targeted the pIC (with well-established roles in both processing and regulating cardiac sensory signals and anxiety-related behaviours[45,46,47,48]) and the medial prefrontal cortex (mPFC) (crucially involved in reward and aversion processing, as well as associated with cardiovascular arousal[49]). To perform simultaneous optogenetic inhibition of the cortex and optically paced tachycardia, we bilaterally injected AAVdj-hSyn::iC++-eYFP or AAVdj-hSyn::eYFP control virus and implanted fibre-optic cannulas into the pIC or mPFC of mice expressing ChRmine in the heart[50] (Fig. 4a). As expected, in eYFP-expressing control mice, we found that 473-nm illumination of the pIC did not affect the reduction of water-seeking by optical pacing in trials with a risk of shock (Fig. 4b,c). By contrast, the same intervention in mice expressing iC++ instead of eYFP alone in pIC did reverse the reduction of water-seeking (Fig. 4d–g); all mice receiving pIC inhibition completed the water-retrieving task and exhibited decreased apprehension, with the time to next lever press after shock reduced to near baseline levels (Fig. 4g). Similarly, iC++ inhibition

increased open-arm exploration time during the EPM assay in optically paced mice relative to eYFP controls (Fig. 4h).

Fig. 4: Optogenetic inhibition of the posterior insula attenuates the anxiogenic response from optical pacing.



a, Illustration of the experimental protocol for simultaneous optically induced tachycardia and optogenetic inhibition of the pIC or mPFC using AAVdj-hSyn::iC++-YFP or control virus (YFP only). **b**, Left, conditions for the Vogel conflict task, in which mice received both 473-nm constant illumination in the pIC or mPFC and optically induced tachycardia during the behavioural task. Right, illustration of the experimental protocol for simultaneous optogenetic inhibition of the pIC with optical pacing during the EPM test. **c**,**d**, Cumulative lever presses during baseline (day 1) and 10% shock (day 2) sessions for mice expressing control (YFP) (**c**) or iC++ (**d**) in the pIC with optical pacing (*n* = 6 mice per group). **e**, Cumulative number of lever presses completed in each session. Owing to increased apprehension, only 1 out of 6 control mice completed the 50-lever-press session on the 10% shock session (*n* = 6 per group; two-sided Wilcoxon rank-sum test, **P*

= 0.0152). **f**, Average lever-pressing rate for 0% and 10% shock experimental sessions. Note that iC++ inhibition partially restores overall rates of lever pressing, but not to baseline levels (*n* = 6 per group; two-way ANOVA with Bonferroni post hoc test: group × condition interaction $F_{(1,10)}$ = 5.533, P = 0.0405; group (opsin) effect $F_{(1,10)}$ = 7.439, P = 0.0213; condition (shock) effect $F_{(1,10)}$ = 67.8, $P < 0.0001$. Bonferroni post hoc: 0% shock YFP versus iC++, $P > 0.9999$; 10% shock YFP versus iC++, **P = 0.0036; YFP 0% versus 10% shock, ****$P < 0.0001$; iC++ 0% versus 10% shock, **P = 0.0039). **g**, Time to next lever press after shock. Note that iC++ inhibition reduces apprehension to no-shock levels (*n* = 30, 17, 30 and 30 presses per group in YFP 0%, YFP 10%, iC++ 0% and iC++ 10%; two-way ANOVA with Bonferroni post hoc test: group (opsin) × condition (shock) interaction $F_{(1,103)}$ = 8.7, P = 0.0039; group (opsin) effect $F_{(1,103)}$ = 8.6, P = 0.0041; condition (shock) effect $F_{(1,103)}$ = 35.6, $P < 0.0001$. Bonferroni post hoc: YFP 0% versus 10%, ****$P < 0.0001$; iC++ 0% versus 10%, P = 0.099; 0% YFP versus iC++, $P > 0.9999$; 10% YFP versus iC++, **P = 0.0011). **h**, Time spent in open arms during 5-min epochs of EPM exploration (*n* = 6 per group; two-way repeated-measures ANOVA with Bonferroni post hoc test: group (opsin) × time interaction $F_{(2,20)}$ = 3.543, P = 0.0482; group (opsin) effect $F_{(1,10)}$ = 1.251, P = 0.2894; time effect $F_{(2,20)}$ = 3.058, P = 0.0694. Bonferroni post hoc: ON epoch YFP versus iC++, *P = 0.0323). **i**,**j**, Cumulative lever presses with the same conditions as **c**,**d** but with expression of YFP (**i**) or iC++ (**j**) in the mPFC (*n* = 6 mice). **k**, Cumulative number of lever presses during each session. Note that in 10% shock sessions, both YFP- and iC++-expressing mPFC mice cease lever pressing (*n* = 6 per group; two-sided Wilcoxon rank-sum test, P = 0.2987). **l**, Average lever-pressing rate for 0% and 10% shock sessions (*n* = 6 per group; two-way repeated-measures ANOVA with Bonferroni post hoc test: group × condition interaction $F_{(1,10)}$ = 0.002521, P = 0.9609; group (opsin) effect $F_{(1,10)}$ = 0.4370, P = 0.5235; condition (shock) effect $F_{(1,10)}$ = 154.1, $P <$ 0.0001. Bonferroni post hoc: 0% shock YFP versus iC++, $P > 0.9999$; 10% shock YFP versus iC++, $P > 0.9999$; YFP 0% versus 10% shock, ****$P = 1.1 \times 10^{-5}$; iC++ 0% versus 10% shock, ****$P = 1.0 \times 10^{-6}$). **m**, Time to next lever press after shock (*n* = 30, 22, 30 and 19 presses per group in mPFC YFP 0%, YFP 10%, iC++ 0% and iC++ 10%; two-way ANOVA with Bonferroni post hoc test: group (opsin) × condition (shock) interaction

$F_{(1,97)} = 3.703$, $P = 0.05725$; group (opsin) effect $F_{(1,97)} = 3.610$, $P = 0.0604$; condition (shock) effect $F_{(1,97)} = 54.18$, $P < 0.000001$. Bonferroni post hoc: YFP 0% versus 10%, ***$P = 0.0009$; iC++ 0% versus 10%, ****$P = 2.95 \times 10^{-8}$; 0% YFP versus iC++, $P > 0.9999$; 10% YFP versus iC++, $P = 0.0698$). **n**, Time spent in open arms during 5-min epochs of EPM exploration with (iC++, blue) and without (YFP, grey) mPFC inhibition ($n = 6$ per group; two-way repeated-measures ANOVA with Bonferroni post hoc test: group (opsin) × time interaction $F_{(2,20)} = 0.3929$, $P = 0.6802$; group (opsin) effect $F_{(1,10)} = 0.00039$, $P = 0.9846$; time effect $F_{(2,20)} = 17.41$, $P < 0.0001$. Bonferroni post hoc: ON epoch YFP versus iC++, $P > 0.9999$).

The attenuation of the anxiogenic effect of optical pacing exhibited specificity to pIC inhibition; inhibition of the mPFC did not decrease cardiac-associated anxiogenic behaviours relative to eYFP controls (Fig. 4i–n). To test for any direct anxiolysis from optogenetic inhibition of the pIC (that is, not through modulation of the cardiac-pacing effect), we performed the EPM assay and the lever-pressing task in the absence of cardiac pacing and with 30% shock trials to allow the detection of pacing-independent apprehensive behaviour (Extended Data Fig. 9a–d). Inhibition of the pIC without cardiac pacing did not increase open-arm exploration, affect lever-press suppression or influence heart rate (Extended Data Fig. 9). Thus, pIC inhibition alone appeared to be insufficient to induce anxiolysis, consistent with previous reports[48]. Together, these results are in line with a model in which the pIC is important for mediating the anxiety-related and apprehensive behaviours that arise from direct cardiac pacing.

Discussion

In this study, we have developed a method for noninvasive optogenetic control of specific cardiac rhythms during active behaviour. We show that the optically induced tachycardia was not intrinsically aversive, but rather elicited anxiety-like behaviours and apprehension in potentially risky environments. Although diverse mechanisms may contribute to this effect, we consider that anxiogenic effects of evoked tachycardia are not likely to be mediated through a reduction in blood pressure[51], as drugs that reduce systolic blood pressure tend to be anxiolytic (for example, propranolol and

clonidine) or neutral (for example, Ca^{2+}-channel blockers). Our observation of anxiogenesis in response to increased heart rates (900 bpm or 15 Hz) is in line with clinical observations that accelerated heart rates—but not other forms of altered haemodynamics (for example, increased heart rate variability)—are associated with panic and other anxiety-related disorders[52,53]. The altered rate, rather than the external nature of cardiac contraction timing, appears to be important; for example, we found that intermittent or asynchronous stimulation close to baseline heart rates at 660 bpm (11 Hz) did not result in anxiety-like behaviour.

In further investigations of the mechanisms that underlie these behaviours, we found that optogenetic pacing activated the pIC, consistent with studies of cardiovascular control in anaesthetized rodents[54] and neuroimaging studies of cardiac interoception and reflex control in humans[45,46,47], including in the setting of panic and anxiety[49,55]. It remains unclear whether this pathway can be modulated by peripheral baroreceptive sensory neurons or other sensory mechanisms that detect changes in blood pressure[56,57,58], and it is possible that there are additional ways in which cardiac viscerosensory information can be relayed to higher cortical areas[41,42]. The anxiogenic behavioural effects of cardiac pacing were attenuated during optogenetic inhibition of the pIC, suggesting that the insula has a causal role in integrating sensory information from the heart with a contextual assessment of environmental risk to produce adaptive behavioural patterns. Our findings support the idea that the insular cortex is involved in monitoring not only consummatory[59,60,61] but also entirely internal interoceptive states to instruct relevant behavioural responses, as predicted from human neuroimaging studies of cardiac interoception[45,46,47].

This study shows that cell-type-specific, temporally precise, noninvasive perturbation of organ-scale physiology is possible in fully intact, freely behaving mammals. Although we have applied our approach mainly to study animal behaviour over a period of minutes, future integration with miniaturized wireless devices[15] may facilitate longer-term studies to modulate targeted populations of cells over days to weeks while alleviating the need for intimate contact with the light source. Furthermore, refinement in cell-type-targeting strategies may enable minimally invasive to noninvasive optogenetic dissection of specific cell types (for example,

pacemaker cells[28], Purkinje fibres[27] and cardiac ganglions[12]) to determine their effects on regulating cardiac electrophysiology and behaviour. Finally, our approach, which requires no specialized optoelectronics or surgery, has the potential for broad application to a range of physiological systems throughout the body—opening up numerous opportunities to explore the complex interactions between physiological systems in health, disease and treatment.

Methods

Mice

All animal procedures followed animal care guidelines approved by Stanford University's Administrative Panel on Laboratory Animal Care (APLAC) and guidelines of the National Institutes of Health. Investigators were not blinded to the genotypes of the mice. Male and female wild-type C57BL6/J (JAX 0064) mice were used for most behavioural experiments unless specified otherwise, and all mice were 8–12 weeks old at the time of starting behavioural experiments. Mice were housed in plastic cages with disposable bedding on a standard light cycle with food and water available ad libitum, except when placed on water restriction. When on water restriction, mice were provided with 1 ml of water each day and maintained above 85% of baseline weight. Behavioural experiments were performed during the dark phase.

Molecular cloning

A 685-bp fragment containing the promoter region of the mouse troponin gene was amplified from a wild-type mouse using CGCACGCGTGAGGCCATTTGGCTCATGAGAAGC and CATGGATCCTCTAGAAAGGGCCATGGATTTCCTG primers, cloned upstream of ChRmine-p2A-oScarlet using MluI and BamHI sites in an AAV backbone, sequence-verified and tested for expression in dissociated neonatal cardiomyocytes.

In vitro cardiomyocyte experiments

Dissociated neonatal mouse cardiomyocytes prepared using the Pierce Isolation Kit (Thermo Fisher Scientific, 88289) were transfected with rAAV-mTNT::ChRmine-p2A-oScarlet (1 µl of 8 × 12 viral genomes (vg) per ml in 500 µl of medium). Three to five days after infection, individual cardiomyocytes were identified under a light microscope. Optical stimulation was provided by a Spectra X Light engine at 585 nm (LumenCore) coupled to a Leica DM LFSA microscope and synchronized with video recording at 100 fps using LabView software. Laser power leaving the imaging objective was measured with an optical power meter (Thorlabs PM100D). Videos were analysed for contraction using custom scripts in MATLAB.

In vivo systemic viral delivery

Wild-type mice aged three to four weeks were anaesthetized with isoflurane and rAAV-mTNT::ChRmine-p2A-oScarlet (2×10^{11} vg per mouse) or vehicle was delivered by retro-orbital injection. Our selected titres were previously used for systemic viral transduction of ChR2 in the heart[26]. A total volume of 60 µl 0.9% NaCl saline solution was injected into the right retro-orbital sinus using a 28G needle, then allowed to recover on a warming pad before being returned to the home cage.

Optical pacemaker in vivo characterization

Mice were tested three weeks after injection of the pacemaker virus. Electrocardiography signals were collected using commercial instruments (Rodent Surgical Monitor+, Indus Instruments), with anaesthetized mice placed in a supine position and limbs placed in contact with electrode pads via a conductive gel. A 594-nm laser (LaserGlow) was attached to a fibre-optic patch cord (Thorlabs) terminating in a 200-µm-diameter, 0.39-NA fibre (Thorlabs) which was positioned against the chest. Optical power was adjusted using the laser's built-in power modulator and measured with an optical power meter (Thorlabs) at the fibre tip. Stimulation was performed with a pulse width of 10 ms and an inter-pulse interval ranging from 120 ms (equivalent to 500 bpm) to 67 ms (900 bpm), controlled by a TTL signal generator (Master-8). Heart rate (bpm) was derived from the heart rate interval between successive R waves (RR interval) obtained from ECG

recordings. Fidelity of photoactivated QRS complexes was quantified by counting the number of beats at a set frequency divided by the number of total beats measured during the middle 20 s of a 30-s stimulation period.

Measurements of systolic blood pressure

Mice were anaesthetized (1.5–2% isofluorane) and placed in the supine position with the chest shaved. Systolic blood pressure measurements were performed using a 1.4-F pressure sensor mounted Millar catheter (SPR-671, ADInstruments) and recorded using LabChart 7 Pro (ADInstruments). The catheter was inserted via the right carotid artery into the left ventricle. A 589-nm laser was used to deliver 240 mW mm^{-2} light across intact chest at either constant or intermittent (500 ms ON, 1,500 ms OFF) optical stimulation at 900 bpm with a 10-ms pulse width for 30 s to assess optogenetic pacing effects on systolic blood pressure in real time.

Wearable optical pacemaker hardware

Custom-made wearable optical stimulators were constructed using 3 × 4.5 mm 591-nm PC Amber Rebel LEDs (Luxeon LXM2-PL01-0000). 30AWG flexible silicone wire (Striveday) was soldered to the LED pad and coated with electrically insulating, thermally conductive epoxy (Arctic Alumina), and adhered to copper sheet cut to 10 × 15 mm for thermal dissipation and subsequently glued to a fabric vest designed for freely moving mouse behaviour (Coulbourn A71-21M25). Wiring was held in place on the vest using hot glue and the free ends were inserted into a breadboard for stimulus control by an LED Driver (Thorlabs LEDD1B T-Cube). The optical power was set to 160–240 mW mm^{-2} measured from the surface of the LED. Light was delivered at intervals consisting of a 10-ms pulse width at 15 Hz (900 bpm) for 500 ms with 1,500 ms OFF time by using either a Master-8 or an Arduino microcontroller synchronized to behaviour recording software. Computer-aided design schematics were created with Onshape. Thermal measurements were performed using a FLIR C2 Compact thermal camera (FLIR) and the thermal profile at the surface of the micro-LED is plotted in Extended Data Fig. 4c.

Freely moving behaviour with pacemaker

All mice were habituated to the experimenter and handled for at least three days, and in addition allowed to acclimatize to wearing optical pacemaker hardware for at least five days, before behavioural experiments. Fur over the chest was removed (Nair) at least five days before behavioural experiments. Mice were briefly anaesthetized with isoflurane before the placement of the optical pacing vest and allowed to fully recover in the home cage (at least 1 h) before experiments. We used a stimulation protocol consisting of a 10-ms pulse width at 15 Hz (900 bpm) with 500 ms ON time and 1,500 ms OFF time to introduce intermittent tachycardia or 10-ms pulse width at 11 Hz (660 bpm) with a Poisson distribution to introduce increased heart rate variability. Mice received optical stimulation during the ON periods of the behavioural assay from the wearable micro-LED device in both control and ChRmine cohorts. No statistical difference in behaviour was observed between virally transduced and control groups at baseline, suggesting that there were no side effects from transgene delivery. No statistical difference in behaviour was observed in control groups before, during and after optical stimulation, suggesting that there were no effects from light delivery alone.

RTPP

Mice were placed in a custom-built RTPP chamber (30.5 × 70 cm) on day 1 to determine their baseline preference for each side of the chamber. Behavioural tracking was performed using blinded automated software (Noldus Ethovision). On day 2, mice were stimulated whenever they were on one side of the chamber. Stimulation sides were randomly assigned and counterbalanced across mice. Each session lasted 20 min.

EPM

The EPM was made of grey plastic (Med Associates). Mice were gently placed in the closed arm of the EPM. Mice were allowed to freely explore the maze for a 5-min baseline ‘off’ period, followed by a 5-min ‘on’ period during which optical stimulation was delivered, and finally a 5-min ‘off’ period. Behavioural tracking was performed using blinded automated software (Noldus Ethovision) and the overall time spent in open arms was reported for each epoch.

OFT

Mice were placed in a 60 × 60-cm arena and allowed to freely explore during a 9-min session. Optical stimulation was delivered during the middle 3-min epoch. Movement was tracked with a video camera positioned above the arena. To assess anxiety-related behaviour, the chamber was divided into a peripheral and centre (48 × 48 cm) region.

Operant lever-pressing task

Water-restricted mice were trained to lever press for a small water reward (around 10 μl water) while freely moving in an operant condition box containing a single retractable lever and a shock grid floor (Coulbourn). Mice were allowed to retrieve a maximum of 50 rewards per day, and sessions were terminated after all rewards had been retrieved or after 30 min. After each lever press, the lever was retracted for 5 s before extending again. After mice retrieved 50 rewards for at least 3 consecutive days (typically 2–3 weeks of training), they were allowed to proceed with stimulation experiments. On shock days, mice were given a 1-s, 0.1-mA foot shock after 10% of lever presses instead of water. Shocks were delivered in a pseudorandom order on lever-press trials 5, 13, 24, 31 and 44, and the time to the next lever press was measured from the time elapsed for these trials until the subsequent lever press. During stimulation experiments (both baseline and shock days), optical stimulation was delivered throughout the experiment. Water was delivered using a custom set-up consisting of a lick spout (Popper and Sons, stainless steel 18-gauge) and a solenoid (Valcor, SV74P61T1) controlled by a microcontroller (Arduino Uno R3). Licking was monitored using a capacitive sensing board (Arduino Tinker Kit) wired to the lick spout and interfacing with the microcontroller. Shocks were delivered using an 8-pole scrambled shock floor (Coulbourn). Behavioural stimuli—lever presentations and retractions, and shocks—were controlled with Coulbourn Graphic State software. The timing of lever presses and licks was also recorded at 5 kHz using a data-acquisition hardware (National Instruments, NI PCIe-6343-X).

TRAP2 labelling

Fos$^{2A\text{-}iCreER}$ (TRAP2; JAX 030323) mice were backcrossed onto a C57BL6/J background and bred with *B6;129S6-Gt(ROSA)26Sor*$^{tm14(CAG\text{-}tdTomato)/Hze}$*/J* (Ai14; JAX 007908) mice, as previously described[38]. Both male and female mice were used for TRAP2 labelling experiments. Mice were injected retro-orbitally with rAAV9-mTNT-ChRmine-oScarlet or a vehicle control at three to four weeks of age. Four weeks later, mice were handled and acclimatized to fresh clean cages and optical pacing equipment for at least seven days before labelling. On the day of labelling, mice were allowed to acclimatize to optical pacing equipment for at least 2 h in a fresh clean cage with food and water, stimulated for 15 min (10-ms pulse width at 15 Hz for 500 ms every 1,500 ms) and left undisturbed for 2 h, at which time they were injected intraperitoneally with 5 mg kg^{-1} 4-hydroxytamoxifen (Sigma) dissolved in normal saline containing 1% Tween-80 and 2.5% DMSO (as described previously[37,39]). Mice were then returned to their home cage and were euthanized at least two weeks later to allow for full expression of the fluorophore.

Whole-brain CLARITY and analysis

Mice were perfused with ice-cold phosphate-buffered saline (PBS) and 4% paraformaldehyde (PFA), then post-fixed in a 1% CLARITY hydrogel solution (1% acrylamide, 0.003125% bis-acrylamide, 4% PFA and 0.25% VA-044 in 1× PBS) for 2 days. Tissue was degassed, polymerized at 37 °C for 4 h and washed with 200 mM sodium borate with 4% sodium dodecyl sulfate solution overnight. Tissue was then electrophoretically cleared for 3–7 days at 80 V (Life Canvas), passively cleared for an additional 2 days, then washed in PBS containing 0.2% Triton-X and 0.02% sodium azide at least 6 times at 37 °C. Cleared samples were refractive-index-matched using RapiClear (Sunjin Labs) and imaged on a custom-built light-sheet microscope[62] using a 10× objective and 5-μm step size or an LaVision Ultramicroscope with a 0.63× zoom macro lens with a step size of 5 μm. Images were visualized using Vision4D (Arivis).

For automated whole-brain registration and cell-segmentation analysis, images were loaded onto Arivis Vision4D software, and neurons were segmented using a built-in supervised pixel-based classifier package based on Ilastik[63] ('Trainable Segmenter'). Segmentation masks were converted to

binary cell masks. Raw light-sheet microscopes images and cell masks were registered to a common reference space defined by the Allen Institute's Reference Atlas and analysed in a region-based manner using our MIRACL package[40].

Induction of *Fos* after pacing

Mice were injected retro-orbitally with AAV9-mTNT-ChRmine-oScarlet or vehicle at three to four weeks of age. At four weeks after injection, mice were handled and acclimatized to fresh clean cages and optical-pacing equipment for a minimum of seven days before pacing experiments. On the day of labelling, mice were allowed to acclimatize to optical-pacing equipment for at least 2 h in a fresh clean cage with food and water, stimulated for 15 min and euthanized 30 min after stimulation by perfusion with ice-cold PBS and 4% PFA under heavy anaesthesia. Tissue was post-fixed in 4% PFA on ice for an additional 24 h (brain) before staining and imaging.

In situ hybridization

Post-fixed brains were cut with a vibratome into 65-µm coronal slices. Heart and other organs were sliced at 200-µm thickness. Tissue slices were stored in 70% ethanol at −20 °C. Established protocols for third-generation hairpin chain reaction (HCR) in situ hybridization were used for coronal slice[64]. In situ hybridization probes (ChRmine, *Fos* and *Slc6a2*) were designed by and purchased from Molecular Instruments. Hybridization was performed overnight in hybridization buffer (Molecular Instruments) at 4 nM probe concentration. The next day, slices were washed (three times in wash buffer at 37 °C then twice in 2× SSCT at room temperature; 30 min each) and then incubated in amplification buffer. Dye-conjugated hairpins (B1-647, B3-488 and B5-546) were heated to 95 °C for 1 min and then cooled to 4 °C. Hairpin amplification was performed by incubating individual slices in 50 µl of amplification buffer with B1, B3 and B5 probes at concentrations of 240 nM overnight in the dark. Samples were stained with DAPI, washed three times with 5× SSCT for 30 min each and then equilibrated in exPROTOS (125 g iohexol, 3 g diatrizoic acid and 5 g *N*-methyl-d-glucamine dissolved in 100 ml deionized water with the refractive index adjusted to 1.458) (ref. [65]), a

high-refractive-index mounting solution, then imaged. Slices were imaged on a confocal microscope (Olympus FV3000).

Cardiac histology

At 48 h post-fixation, hearts were sectioned into 200-μm slices. For staining, slices were first incubated for 10 min in blocking solution (3% normal donkey serum (NDS) in PBST), followed by primary antibody staining overnight at 4 °C using the following antibodies: anti-vimentin (ab24525), anti-cardiac troponin I (ab188877) or anti-PGP9.5 (ab108986), purchased from Abcam at 1:200 dilution in blocking solution. Slices were then washed twice in PBST, then stained with secondary antibodies (1 mg ml^{-1}) at 1:500 dilution for 3 h at room temperature using the following: F(ab')2 anti-chicken 488 (703-546-155) and anti-rabbit 647 (711-606-152) purchased from Jackson ImmunoResearch Laboratories. The slices were then stained with DAPI and washed three times with PBST (30 min per wash). Sections were mounted onto slides and mounted with exPROTOS. Slices were imaged on a confocal microscope (Olympus FV3000).

Stereotaxic surgery for optogenetic experiments

For all surgeries, mice were anesthetized with 1–2% isoflurane, and placed in a stereotaxic apparatus (Kopf Instruments) on a heating pad (Harvard Apparatus). Fur was removed from the scalp, the incision site was cleaned with betadine and a midline incision was made. Sterile surgical techniques were used, and mice were injected with sustained-release buprenorphine for post-operative recovery. Mice were allowed to recover for at least two weeks after surgery before behavioural experiments.

For intracranial optogenetic experiments, virus was injected using a 33-gauge beveled needle and a 10-μl Nano-fil syringe (World Precision Instruments), controlled by an injection pump (Harvard Apparatus). Five hundred nanolitres of AAVdj-hSyn::iC++-eYFP or AAVdj-hSyn::eYFP (5×10^{11} vg ml^{-1}) was injected at 150 nl per min and the syringe was left in place for at least 10 min before removal. The following coordinates were used (relative to Bregma): posterior insula (−0.58 (anterior–posterior (AP)), ±4.2 (medial–lateral (ML)), −3.85 (dorsal–ventral (DV)); mPFC (1.8 (AP),

±0.35 (ML), −2.9 (DV)). Optical fibres (0.39 NA, 200 μm; Thorlabs) were implanted 200 μm above virus injection coordinates. Fibres were secured to the cranium using Metabond (Parkell). Mice were allowed to recover for at least two weeks before behavioural testing.

In vivo electrophysiology

The mice with or without cardiac-targeted ChRmine expression were implanted with custom-made headplates, reference electrodes and cyanoacrylate-adhesive-based 'clear-skull caps' as previously described[66]. After recovery, mice were water-restricted and habituated to head fixation, but they were allowed to drink water to satiate thirst before recording sessions. Craniotomies were made with a dental drill at least several hours before recording sessions and were sealed with Kwik-Cast (World Precision Instruments). Exposed craniotomies before, during and after recordings were kept moist with frequent application of saline until sealed with Kwik-Cast.

Before recordings, the mice were placed into the pacemaker vests and reliable pacing was confirmed by ECG under brief anaesthesia with isoflurane. Then the mice were head-fixed and allowed to recover. Next, one or two (for simultaneous bilateral recordings) four-shank Neuropixels 2.0 probes mounted on a multi-probe manipulator system (New Scale Technologies) and controlled by SpikeGLX software (Janelia Research Campus) were inserted through the craniotomies at variable angles (0–20°) depending on the recording geometry. Typically the probes were aimed to touch the skull around the insula, which could be inferred from probe bending or changes in local field potential, and then were retracted around 100 μm and allowed to sit in place for at least 15 min before recordings. Recordings were performed along each of the four shanks sequentially while mice received 5 s of optical stimulation (900 bpm (15 Hz)) with inter-trial intervals of at least 15–25 s. Probes were cleaned with trypsin between recording sessions. Spike sorting was performed by Kilosort 2.5 and auxiliary software as previously described[66].

After recordings, the brains were perfused, cleared, imaged and registered to the Allen Brain Atlas as previously described[66]. Using the traces of lipophilic dye CM-DiI or DiD (which coated the probes before each

insertion) and electrophysiological features, the atlas coordinates of the recorded single units were determined.

The spikes from single units were aligned to pacing onset, and the visualized peri-stimulus time histograms were calculated by subtracting 5 s baseline firing rate, 10 ms binning and 500 ms half-Gaussian filtering. The population-averaged firing rate of each region was calculated by combining *z*-scores (before filtering) over time for all single units in the region of interest. Specifically, we used hierarchical bootstrap to combine data from multiple levels as previously described[66]. For each condition, 100 bootstrap datasets were generated, and their mean and s.d. represented the mean and s.e.m. of the initial dataset. For statistical tests comparing ChRmine and control groups, the one-sided *P* value for the null hypothesis (the ChRmine firing rate subtracted by the control firing rate is zero) was calculated as the fraction of these subtracted values from the pairs of the resampled means (averaged over the time window of interest) that were smaller than zero.

Optogenetic freely moving behaviour

For optogenetic inhibition of iC++, a 473-nm laser (Omicron Laserage) was used to deliver constant light at 2–3 mW measured from the tip. Laser shutters were controlled using a Master-8 receiving synchronized input from behaviour apparatus and control software (Ethovision).

Statistical analysis

The target number of subjects used in each experiment was determined on the basis of numbers in previously published studies. No statistical methods were used to predetermine sample size or randomize. Criteria for excluding mice from analysis are listed in the methods. Mean ± s.e.m. was used to report statistics. The statistical test used, definition of *n* and multiple-hypothesis correction where appropriate are described in the figure legends. Unless otherwise stated, all statistical tests were two-sided. Significance was defined as alpha = 0.05. All statistical analyses were performed in GraphPad Prism 9.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All primary data for all figures and extended data figures are available from the corresponding author upon request.

References

1. Critchley, H. D. & Harrison, N. A. Visceral influences on brain and behavior. *Neuron* **77**, 624–638 (2013).

2. Chen, W. G. et al. The emerging science of interoception: sensing, integrating, interpreting, and regulating signals within the self. *Trends Neurosci.* **44**, 3–16 (2021).

3. Craig, A. D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* **3**, 655–666 (2002).

4. Critchley, H. D. & Garfinkel, S. N. Interoception and emotion. *Curr. Opin. Psychol.* **17**, 7–14 (2017).

5. Tural, U. & Iosifescu, D. V. The prevalence of mitral valve prolapse in panic disorder: a meta-analysis. *Psychosomatics* **60**, 393–401 (2019).

6. Frommeyer, G., Eckardt, L. & Breithardt, G. Panic attacks and supraventricular tachycardias: the chicken or the egg? *Neth. Heart J.* **21**, 74–77 (2013).

7. Oppenheimer, S. & Cechetto, D. The insular cortex and the regulation of cardiac function. *Compr. Physiol.* **6**, 1081–1133 (2011).

8. Schandry, R. Heart beat perception and emotional experience. *Psychophysiology* **18**, 483–488 (1981).

9. Lange, C. G. & James, W. *The Emotions* (Williams & Wilkins, 1922).

10. Gogolla, N. The insular cortex. *Curr. Biol.* **27**, R580–R586 (2017).

11. Klein, A. S., Dolensek, N., Weiand, C. & Gogolla, N. Fear balance is maintained by bodily feedback to the insular cortex in mice. *Science* **374**, 1010–1015 (2021).

12. Rajendran, P. S. et al. Identification of peripheral neural circuits that regulate heart rate using optogenetic and viral vector strategies. *Nat. Commun.* **10**, 1944 (2019).

13. Ressler, K. J. & Mayberg, H. S. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nat. Neurosci.* **10**, 1116–1124 (2007).

14. Entcheva, E. & Kay, M. W. Cardiac optogenetics: a decade of enlightenment. *Nat. Rev. Cardiol.* **18**, 349–367 (2021).

15. Won, S. M., Song, E., Reeder, J. T. & Rogers, J. A. Emerging modalities and implantable technologies for neuromodulation. *Cell* **181**, 115–135 (2020).

16. Shahriari, D., Rosenfeld, D. & Anikeeva, P. Emerging frontier of peripheral nerve and organ interfaces. *Neuron* **108**, 270–285 (2020).

17. Montgomery, K. L., Iyer, S. M., Christensen, A. J., Deisseroth, K. & Delp, S. L. Beyond the brain: optogenetic control in the spinal cord and peripheral nervous system. *Sci. Transl. Med.* **8**, 337rv5 (2016).

18. Berntson, G. G. & Khalsa, S. S. Neural circuits of interoception. *Trends Neurosci.* **44**, 17–28 (2021).

19. Mickle, A. D. et al. A wireless closed-loop system for optogenetic peripheral neuromodulation. *Nature* **565**, 361–365 (2019).

20. Hulsmans, M. et al. A miniaturized, programmable pacemaker for long-term studies in the mouse. *Circ. Res.* **123**, 1208–1219 (2018).

21. Bilchick, K. C. et al. Differential regional gene expression from cardiac dyssynchrony induced by chronic right ventricular free wall pacing in the mouse. *Physiolog. Genomics* **26**, 109–115 (2006).

22. Laughner, J. I. et al. A fully implantable pacemaker for the mouse: from battery to wireless power. *PLoS ONE* **8**, e76291 (2013).

23. Gutruf, P. et al. Wireless, battery-free, fully implantable multimodal and multisite pacemakers for applications in small animal models. *Nat. Commun.* **10**, 5742 (2019).

24. Nussinovitch, U. & Gepstein, L. Optogenetics for in vivo cardiac pacing and resynchronization therapies. *Nat. Biotechnol.* **33**, 750–754 (2015).

25. Nyns, E. C. A. et al. Optogenetic termination of ventricular arrhythmias in the whole heart: towards biological cardiac rhythm management. *Eur. Heart J.* **38**, 2132–2136 (2016).

26. Vogt, C. C. et al. Systemic gene transfer enables optogenetic pacing of mouse hearts. *Cardiovasc. Res.* **106**, 338–343 (2015).

27. Zaglia, T. et al. Optogenetic determination of the myocardial requirements for extrasystoles by cell type-specific targeting of ChannelRhodopsin-2. *Proc. Natl Acad. Sci. USA* **112**, E4495–E4504 (2015).

28. Lee, F. K. et al. Genetically engineered mice for combinatorial cardiovascular optobiology. *eLife* **10**, e67858 (2021).

29. Marshel, J. H. et al. Cortical layer-specific critical dynamics triggering perception. *Science* **365**, eaaw5202 (2019).

30. Chen, R. et al. Deep brain optogenetics without intracranial surgery. *Nat. Biotechnol.* **39**, 161–164 (2021).

31. Pacak, C. A. et al. Recombinant adeno-associated virus serotype 9 leads to preferential cardiac transduction in vivo. *Circ. Res.* **99**, e3–e9 (2006).

32. Bedbrook, C. N. et al. Machine learning-guided channelrhodopsin engineering enables minimally invasive optogenetics. *Nat. Methods* **16**, 1176–1184 (2019).

33. Kishi, K. E. et al. Structural basis for channel conduction in the pump-like channelrhodopsin ChRmine. *Cell* **185**, 672–689 (2022).

34. Sgoifo, A., Carnevali, L., Pico Alfonso, M. D. L. A. & Amore, M. Autonomic dysfunction and heart rate variability in depression. *Stress* **18**, 343–352 (2015).

35. Lampert, R. et al. Emotional and physical precipitants of ventricular arrhythmia. *Circulation* **106**, 1800–1805 (2002).

36. Taggart, P., Critchley, H. & Lambiase, P. Heart–brain interactions in cardiac arrhythmia. *Heart* **97**, 698–708 (2011).

37. Kim, C. K. et al. Molecular and circuit-dynamical identification of top-down neural mechanisms for restraint of reward seeking. *Cell* **170**, 1013–1027 (2017).

38. Allen, W. E. et al. Thirst-associated preoptic neurons encode an aversive motivational drive. *Science* **357**, 1149 (2017).

39. Ye, L. et al. Wiring and molecular features of prefrontal ensembles representing distinct experiences. *Cell* **165**, 1776–1788 (2016).

40. Goubran, M. et al. Multimodal image registration and connectivity analysis for integration of connectomic data from microscopy to MRI. *Nat. Commun.* **10**, 5504 (2019).

41. Benarroch, E. E. The central autonomic network: functional organization, dysfunction, and perspective. *Mayo Clin. Proc.* **68**, 988–1001 (1993).

42. Azzalini, D., Rebollo, I. & Tallon-Baudry, C. Visceral signals shape brain dynamics and cognition. *Trends Cogn. Sci.* **23**, 488–509 (2019).

43. Carter, M. E. et al. Tuning arousal with optogenetic modulation of locus coeruleus neurons. *Nat. Neurosci.* **13**, 1526 (2010).

44. Livneh, Y. & Andermann, M. L. Cellular activity in insular cortex across seconds to hours: sensations and predictions of bodily states. *Neuron* **109**, 3576–3593 (2021).

45. Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A. & Dolan, R. J. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* **7**, 189–195 (2004).

46. Khalsa, S. S., Rudrauf, D., Feinstein, J. S. & Tranel, D. The pathways of interoceptive awareness. *Nat. Neurosci.* **12**, 1494–1496 (2009).

47. Hassanpour, M. S. et al. The insular cortex dynamically maps changes in cardiorespiratory interoception. *Neuropsychopharmacology* **43**, 426–434 (2018).

48. Gehrlach, D. A. et al. Aversive state processing in the posterior insular cortex. *Nat. Neurosci.* **22**, 1424–1437 (2019).

49. Pollatos, O., Schandry, R., Auer, D. P. & Kaufmann, C. Brain structures mediating cardiovascular arousal and interoceptive awareness. *Brain Res.* **1141**, 178–187 (2007).

50. Berndt, A. et al. Structural foundations of optogenetics: determinants of channelrhodopsin ion selectivity. *Proc. Natl Acad. Sci. USA* **113**, 822–829 (2016).

51. Carnovale, C. et al. Antihypertensive drugs and brain function: mechanisms underlying therapeutically beneficial and harmful neuropsychiatric effects. *Cardiovasc. Res.* **2022**, cvac110 (2022).

52. Gorman, J. M. & Sloan, R. P. Heart rate variability in depressive and anxiety disorders. *Am. Heart J.* **140**, S77–S83 (2000).

53. Chalmers, J. A., Quintana, D. S., Abbott, M. J.-A. & Kemp, A. H. Anxiety disorders are associated with reduced heart rate variability: a meta-analysis. *Front. Psychiatry* **5**, 80 (2014).

54. Oppenheimer, S. M. & Cechetto, D. F. Cardiac chronotropic organization of the rat insular cortex. *Brain Res.* **533**, 66–72 (1990).

55. Schulz, S. M. Neural correlates of heart-focused interoception: a functional magnetic resonance imaging meta-analysis. *Phil. Trans. R. Soc. B* **371**, 20160018 (2016).

56. Zeng, W.-Z. et al. PIEZOs mediate neuronal sensing of blood pressure and the baroreceptor reflex. *Science* **362**, 464 (2018).

57. Min, S. et al. Arterial baroreceptors sense blood pressure through decorated aortic claws. *Cell Rep.* **29**, 2192–2201 (2019).

58. Zhang, Z. & Oppenheimer, S. M. Characterization, distribution and lateralization of baroreceptor-related neurons in the rat insular cortex. *Brain Res.* **760**, 243–250 (1997).

59. Livneh, Y. et al. Homeostatic circuits selectively gate food cue responses in insular cortex. *Nature* **546**, 611–616 (2017).

60. Kusumoto-Yoshida, I., Liu, H., Chen, B. T., Fontanini, A. & Bonci, A. Central role for the insular cortex in mediating conditioned responses to anticipatory cues. *Proc. Natl Acad. Sci. USA* **112**, 1190 (2015).

61. Wu, Y. et al. The anterior insular cortex unilaterally controls feeding in response to aversive visceral stimuli in mice. *Nat. Commun.* **11**, 640 (2020).

62. Tomer, R., Ye, L., Hsueh, B. & Deisseroth, K. Advanced CLARITY for rapid and high-resolution imaging of intact tissues. *Nat. Protoc.* **9**, 1682–1697 (2014).

63. Berg, S. et al. ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).

64. Choi, H. M. et al. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753 (2018).

65. Park, Y.-G. et al. Protection of tissue physicochemical properties using polyfunctional crosslinkers. *Nat. Biotechnol.* **37**, 73 (2018).

66. Sylwestrak, E. L. et al. Cell-type-specific population dynamics of diverse reward computations. *Cell* **185**, 3568–3587 (2022).

Acknowledgements

We thank all members of the K.D. laboratory, particularly M. Lovett-Barron, C. Bedbrook and F. Gore, for comments. This work was supported by grants from the NIH, NSF, Gatsby, Fresenius, Wiegers, Grosfeld and NOMIS Foundations (to K.D.); the Stanford MSTP and Bio-X (to B.H.); and a K99 NS119784 and NARSAD Young Investigator Grant (to R.C.).

Author information

Author notes

1. These authors contributed equally: Brian Hsueh, Ritchie Chen

Authors and Affiliations

1. Department of Bioengineering, Stanford University, Stanford, CA, USA

   Brian Hsueh, Ritchie Chen, YoungJu Jo, Daniel Tang, Misha Raffiee, Yoon Seok Kim, Masatoshi Inoue, Sawyer Randles, Charu Ramakrishnan, Sneha Patel, Doo Kyung Kim, Tony X. Liu, Soo Hyun Kim, Longzhi Tan, Leili Mortazavi, Arjay Cordero, Jenny Shi, Theodore T. Ho, Ailey Crow, Ai-Chi Wang Yoo, Cephra Raja, Kathryn Evans & Karl Deisseroth

2. Department of Pediatrics, Stanford University, Stanford, CA, USA

   Mingming Zhao & Daniel Bernstein

3. Department of Radiology, Stanford University, Stanford, CA, USA

Michael Zeineh & Maged Goubran

4. Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA

Karl Deisseroth

5. Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA

Karl Deisseroth

Contributions

B.H., R.C. and K.D. designed the project and experiments, and wrote the manuscript. B.H., R.C., D.T., M.R., Y.S.K., M.I., S.R., S.P., D.K.K., S.H.K., L.T., L.M., A. Cordero, J.S., M. Zhao, T.H., A. Crow, A.-C.W.Y., C. Raja, K.E. and D.B. contributed to the collection of data and interpretation of results with input from K.D. Y.J. performed electrophysiology recordings and analysis with assistance from T.X.L. and R.C. C. Ramakrishnan designed the viral constructs. B.H., M.G. and M. Zeineh performed whole-brain image registration. K.D. supervised all aspects of this work.

Corresponding author

Correspondence to Karl Deisseroth.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Anna Beyeler, Tobias Bruegmann, Emilia Entcheva and Scott Russo for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 In vitro characterization of the optical pacemaker.

a, Quantified cardiomyocyte contraction sequence measured by centroid motion (see Supplementary Video 1). Optical stimulation was applied at 5 Hz, with 10-ms pulse width at 585 nm. Scale: 1 s, 1 pixel. **b**, Fidelity of optically induced contractions at different light intensities. n = 20 cells. Data represent mean ± s.e.m.

Extended Data Fig. 2 In vivo characterization of AAV9-mTNT::ChRmine-2A-oScarlet.

a,**b**, Representative confocal images of ChRmine-2A-oScarlet (red)-infected ventricular (**a**) and atrial (**b**) cardiac tissue, co-labelled with troponin (white), vimentin (green), and DAPI (blue). Note ChRmine expression is restricted to troponin$^+$ cardiomyocytes with no off-target labelling in neighbouring vimentin$^+$ fibroblasts. Scale bar=100 µm. **c**, Penetrance of AAV9-mTNT::ChRmine-p2A-oScarlet quantified as percentage of troponin$^+$ cells that express ChRmine-2A-oScarlet (n = 3). **d**, Specificity of AAV9-mTNT::ChRmine-p2A-oScarlet quantified as percentage of ChRmine-2A-oScarlet$^+$ cells that are troponin$^+$ (n = 3). **e**, Quantification of oScarlet expression in ventricular and atrial cardiac tissue and liver as mean fluorescence signal (arbitrary unit) following 1 month post-injection of AAV9-mTNT::ChRmine-p2A-oScarlet (n = 3, one-way repeated-measures ANOVA with Bonferroni post-hoc test: $F(2,4)=18.3$, $p = 0.0097$. Post-hoc:

ventricle vs. atrium p = 0.99, ventricle vs. liver p = 0.014, atrium vs liver p = 0.029). **f**, Representative confocal image from one of two mice depicting lack of neuronal labelling in cardiac ganglia following retro-orbital delivery of AAV9-mTNT::ChRmine-p2A-oScarlet as measured by immunostaining for the neural marker PGP9.5 (cyan). No off-target ChRmine expression was observed in neurites (centre) or in the soma of cardiac ganglia (bottom). Scale bar=100 µm (top), 20 µm (center, bottom). **g**,**h**, Representative confocal images of oScarlet (red), ChRmine mRNA (white) and nuclei (DAPI) nuclei across different organs (**g**) after 9 months of expression from one of two mice. No off-target expression of ChRmine was observed in organs beyond the heart, including throughout the brain (**h**). Scale bar=100 µm. Data represent mean ± s.e.m.

Extended Data Fig. 3 Quantification of cardiac responses to optical pacing in vivo.

a, Representative ECG recording with optical pacing at 400 bpm below (top) or above (bottom) a resting heart rate of 400 bpm. **b**, Percentage of photoactivated heart beats that track with the delivered 400 bpm optical stimulus below or above a resting heart of 400 bpm (n = 4 (below), n = 5 (above) mice). **c**, Measured heart rate before, during, and after pacing at 400 bpm in heavily anesthetized mice with a resting heart rate below 400 bpm (n = 4 mice, one-way repeated-measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.01, 3.03)}=34.07$, p = 0.0097; individual $F_{(3,6)}=2.4$, p = 0.17. Post-hoc: OFF vs. ON *p = 0.015, OFF vs. OFF **p = 0.002). **d**, Measured heart rate before, during, and after pacing at 400 bpm in lightly anesthetized mice with a resting heart rate above 400 bpm (n = 5 mice, one-way repeated-measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.06, 4.24)}=3.71$, p = 0.12; individual $F_{(4,8)}=5.9$, p = 0.016. Post-hoc: OFF vs. ON p = 0.17, OFF vs. OFF p = 0.72). **e**, Representative ECG recording with optical pacing delivered at 900 bpm with 10-ms pulse width with laser positioned at the centre of the chest to induce right ventricular pacing (top) with additional traces from before, during, and after light stimulation (bottom). **f**, Representative ECG recording with optical pacing delivered at 900 bpm with 10-ms pulse width with laser positioned more lateral of the chest to induce left ventricular pacing (top) with additional traces from before, during, and after light stimulation (bottom). **g**, Representative QRS

complexes averaged over 100 heart beats before, during, and after pacing for right and left ventricular stimulation. Dark grey bar indicates duration of QRS complex. **h**, QRS duration before, during, and after right ventricular pacing (n = 4 mice, one-way repeated-measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.01, 3.04)}$=25.4, p = 0.015; individual $F_{(3,6)}$=0.68, p = 0.59. Post-hoc: OFF vs. ON *p = 0.037, OFF vs. OFF p = 0.99). **i**, QRS duration before, during, and after left ventricular pacing (n = 4 mice, one-way repeated-measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.18, 3.54)}$=121.7, p = 6.7e-4; individual $F_{(3,6)}$=16.64, p = 0.0026. Post-hoc: OFF vs. ON **p = 0.0038, OFF vs. OFF p = 0.20). **j**, Representative left ventricular blood pressure recordings with sustained optical pacing delivered at 900 bpm with 10-ms pulse width (top) with additional traces from before and after light stimulation (bottom). **k**, Systolic blood pressure (SBP) over time with sustained 900 bpm optical pacing (orange) for 30 s showing a sustained drop in SBP during stimulation (n = 3 mice). **l**, Representative left ventricular blood pressure recordings with intermittent optical pacing delivered at 900 bpm with 10-ms pulse width for 500 ms every 1,500 ms (top) with additional traces during light stimulation (bottom). **m**, Averaged SBP over the interval of optical stimulation (500 ms ON at 900 bpm, 1,500 ms OFF) (n = 3 mice). Data represent mean ± s.d.

Extended Data Fig. 4 Characterization of the wearable micro-LED vest.

a, Schematic of optical pacing vest. **b**, Photographs of freely moving mouse wearing optical pacing vest while receiving 591 nm light stimulation. **c**, Heating of device operated at 900 bpm with 10-ms pulse width at varying duty cycles (n = 3 devices, 25% (500 ms ON, 1,500 ms OFF), 50% (1 s ON, 1 s OFF), 100% (constant ON)). **d**, Percentage of photoactivated QRS complexes as a function of irradiance using the micro-LED device (n = 5). **e**, Measured heart rate across time over a 10 min stimulation period (900 bpm) showing ability of optical pacing to induce a stable heart rhythm (n = 6 mice). **f**, Representative ECG recording from one mouse receiving 900-bpm stimulation for 10 min (top) with additional traces from before, during, and after light stimulation (bottom). **g**, Representative QRS complexes averaged over 100 heart beats before, during and after pacing. **h**, QRS duration before, during and after right ventricular pacing (n = 5 mice, one-way repeated-

measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.03, 4.10)}$=13.74, p = 0.02; individual $F_{(4,8)}$=1.83, p = 0.21. Post-hoc: OFF vs. ON *p = 0.024, OFF vs. OFF p = 0.88). Data represent mean ± s.e.m.

Extended Data Fig. 5 Additional characterization of optical-pacing effects on mouse behaviour.

a–**c**, A hot-plate test was performed to assess for potential effects on thermal pain thresholds from optical pacing (n = 17 (control), 16 (ChRmine)) and the following were quantified: time to first rear (unpaired two-tailed t-test, p = 0.53) (**a**); rears per minute (unpaired two-tailed t-test, p = 0.64) (**b**); and time to first jump (unpaired two-tailed t-test, p = 0.22) (**c**). Note no statistical significance in thermal thresholds was observed with cardiac pacing. **d**–**f**, To assess behavioural differences between control and virally transduced ChRmine-expressing mice, the following comparisons were performed: time spent in one chamber during baseline day (no light delivery) (n = 16 per group, unpaired two-tailed t-test p = 0.21) (**d**); time spent in the open arm of the EPM test during the first 5-min epoch with no light delivery (n = 16 per group, unpaired two-tailed t-test p = 0.61) (**e**); and time spent in the centre of the OFT during the first 3-min epoch with no light delivery (n = 5 (control), 9 (ChRmine), unpaired two-tailed t-test p = 0.15) (**f**). Note no statistical significance in behaviour was observed from viral-transfection. **g**,**h**, To assess for effects of light stimulation alone on mouse behaviour, the following comparisons were performed: time spent in the open arms by control (saline injected) mice during the 15 min EPM assay, where intermittent light delivery (10-ms pulse width, 900 bpm for 500 ms every 2 s) was delivered during the 5 min ON epoch (n = 16, one-way repeated-measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.44,21.62)}$=2.942, p = 0.088, individual $F_{(15,30)}$=1.61, p = 0.13. Post-hoc: OFF vs. ON p = 0.99, OFF vs. OFF p = 0.9, ON vs. OFF p = 0.27) (**g**); and time spent in the centre by control mice during the 9 min OFT, where intermittent light delivery was delivered during the 3 min ON epoch (n = 5, one-way repeated-measures ANOVA with Bonferroni post-hoc test: condition $F_{(1.84,7.45)}$=1.67, p = 0.25, individual $F_{(4,8)}$=1.51, p = 0.29. Post-hoc: OFF vs. ON p = 0.99, OFF vs. OFF p = 0.99, ON vs. OFF p = 0.41) (**h**). Note no statistical significance in behaviour was observed from light

stimulation alone. **i**,**j**, To assess for effects from optical pacing (10-ms pulse width, 900 bpm for 500 ms every 2 s) on anxiety-like behaviour in female mice, the following behavioural assays were measured: time spent in the open arms during the EPM (n = 8 (control) and 7 (ChRmine) female mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,26)}$=1.91, p = 0.17, group (opsin) effect $F_{(1,13)}$=5.1, p = 0.042; time effect $F_{(2,26)}$=4.24, p = 0.026. Bonferroni post-hoc: ON epoch ChRmine vs Control *p = 0.033) (**i**); and time spent in the centre during the OFT (n = 7 (control) and 7 (ChRmine) female mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,24)}$=0.37, p = 0.70; group (opsin) effect $F_{(1,12)}$=7.47, p = 0.018; time effect $F_{(2,24)}$=6.9, p = 0.0043. Post-hoc: ON epoch ChRmine vs Control *p = 0.031) (**j**). **k**,**l**, To determine whether cardiac pacing was aversive in female mice, we also measured the percentage of time spent on baseline and stimulation day for control (grey) and ChRmine-expressing (red) mice during the RTPP assay (n = 7 female mice per group, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x treatment interaction $F_{(1,12)}$=0.68, p = 0.42; group (opsin) effect $F_{(1,12)}$=0.11, p = 0.91; treatment effect $F_{(1,12)}$=0.35, p = 0.57. Post-hoc: Baseline vs Stimulation for Control: p = 0.67, ChRmine: p = 0.99) (**k**); and the average velocity on the optically paced side during RTPP (n = 7 female mice per group, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x treatment interaction $F_{(1,12)}$=0.088, p = 0.77; group (opsin) effect $F_{(1,12)}$=3.69, p = 0.079; treatment effect $F_{(1,12)}$=0.12, p = 0.73. Post-hoc: Stimulation day Control vs Chrmine: p = 0.59) (**l**). **m**,**n**, To determine whether increased heart rate variability can affect anxiety-like behaviour, we introduced constant 660 bpm stimulation with a Poisson distribution in mice and measured the time spent in the centre during the OFT (n = 12 (control) and 14 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,48)}$=1.29, p = 0.28; group (opsin) effect $F_{(1,24)}$=5.80, p = 0.024; time effect $F_{(1.36,32.6)}$=0.47, p = 0.55. Post-hoc ChRmine vs Control: OFF p = 0.61, ON p = 0.28, OFF p = 0.12) (**m**); and the time spent in the open arms during the EPM (n = 14 (control) and 14 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group

(opsin) x time interaction $F_{(2,52)}$=1.75, p = 0.18; group (opsin) effect $F_{(1,26)}$=3.3, p = 0.082; time effect $F_{(1.84,47.8)}$=5.157, p = 0.011. Post-hoc ChRmine vs Control: OFF p = 0.25, ON p = 0.20, OFF p = 0.99) (**n**). **o**,**p**, To determine whether intermittent tachycardia at a rhythm below 900 bpm can affect anxiety-like behaviour, we introduced intermittent 660 bpm stimulation (10-ms pulse width, 660 bpm for 500 ms every 2 s) in mice and measured the time spent in the centre during the OFT (n = 6 (control) and 10 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,28)}$=0.13, p = 0.88; group (opsin) effect $F_{(1,14)}$=0.25, p = 0.63; time effect $F_{(2,28)}$=1.1, p = 0.35. Post-hoc ChRmine vs Control: OFF p = 0.99, ON p = 0.99, OFF p = 0.99) (**o**); and the time spent in the open arms during the EPM (n = 6 (control) and 10 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,28)}$=0.52, p = 0.60; group (opsin) effect $F_{(1,14)}$=0.03, p = 0.86; time effect $F_{(2,28)}$=2.23, p = 0.12. Post-hoc ChRmine vs Control: OFF p = 0.99, ON p = 0.99, OFF p = 0.99) (**p**). **q**, Schematic overview of the chronic stimulation experiment. Mice were stimulated with intermittent optical pacing (900 bpm for 500 ms every 2 s) for 1 h every other day for two weeks before performing the OFT and EPM behavioural assays. **r**, Time spent in open arm during the EPM test for control (grey) and ChRmine (red) mice subjected to chronic optical stimulation (n = 6 (control) and 9 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,26)}$=0.87, p = 0.43; group (opsin) effect $F_{(1,13)}$=0.71, p = 0.41; time effect $F_{(1.65,21.4)}$=0.96, p = 0.38. Post-hoc ChRmine vs Control: 0–5 min p = 0.99, 5–10 min p = 0.89, 10–15 min p = 0.84). **s**, Time spent in centre during the OFT test for control (grey) and ChRmine (red) mice subjected to chronic optical stimulation (n = 6 (control) and 9 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,26)}$=0.25, p = 0.78; group (opsin) effect $F_{(1,13)}$=0.097, p = 0.76; time effect $F_{(1.78,23.2)}$=0.49, p = 0.60. Post-hoc ChRmine vs Control: 0–3 min p = 0.99, 3–6 min p = 0.99, 6–9 min p = 0.99). **t**, Average velocity (cm/s) of mice in the OFT test for control (grey) and ChRmine (red) mice subjected to chronic optical stimulation (n = 6 (control) and 9 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,26)}$=0.29, p = 0.78; group (opsin) effect

$F_{(1,13)}$=0.005, p = 0.94; time effect $F_{(1.34,17.4)}$=17.3, p = 2.6e-4. Post-hoc ChRmine vs Control: 0–3 min p = 0.99, 3–6 min p = 0.99, 6–9 min p = 0.99). **u**, Total distance travelled (cm) mice during the OFT test for control (grey) and ChRmine (red) mice subjected to chronic optical stimulation (n = 6 (control) and 9 (ChRmine) mice, two-way repeated-measures ANOVA with Bonferroni post-hoc test: group (opsin) x time interaction $F_{(2,26)}$=0.29, p = 0.75; group (opsin) effect $F_{(1,13)}$=0.005, p = 0.94; time effect $F_{(1.34,17.4)}$=17.3, p = 2.6e-4. Post-hoc ChRmine vs Control: 0–3 min p = 0.99, 3–6 min p = 0.99, 6–9 min p = 0.99). Data represent mean ± s.e.m.

Extended Data Fig. 6 Operant lever-pressing task at higher risk of shock.

a, Cumulative lever presses during 20% shock session for control (grey) and ChRmine-expressing (red) mice (n = 8 mice per group). Experiment was performed otherwise identically to the behaviour described in Fig. 2. **b**, Lever-pressing rate averaged across the entire 20% shock trial session (n = 8 mice per group, Two-tailed *t*-test, *p = 0.0367). **c**, Cumulative lever presses during 30% shock session. **d**, Lever-pressing rate averaged across the entire 30% shock trial session (n = 8 mice per group, Two-tailed t-test, p = 0.7663). Data represent mean ± s.e.m.

Extended Data Fig. 7 Generation of brain-wide activity maps during optical pacing.

a, Representative whole-brain CLARITY sagittal (top) and axial (bottom) images of control and ChRmine expressing mice exposed to optical pacing. Heart-targeted ChRmine expression resulted in increased tdTomato+ cells throughout the brain (seen as black dots) relative to control. **b**, Following light-sheet imaging, each image stack was registered to a common Allen Reference Atlas using our previously reported computational pipeline[40]. Depicted are coronal slices across the brain with the overlaid anatomical atlas, where different anatomical regions are depicted with different colours. **c**, Representative raw image of single plane of TRAP2-tdTomato brain imaged under light-sheet microscope (left), and after detection by a supervised classifier (Ilastik, Arivis plugin), where single-cells are outlined

in red (right, bottom). Blue arrows indicate example features that were detected by the classifier. **d**, Example CLARITY light-sheet images of control and ChRmine-paced mice in the visceral cortex (VISC), Medulla-behavioural state related (MY), and in the lateral amygdalar nucleus (LA), which are outlined in green in the axial reference slice. Scale bar=500 µm. **e**, Regional cell counts of paced (ChRmine, red) vs control (grey) cohorts sorted from anterior to posterior in all anatomical regions (n = 9 per group, multiple two-sided t-tests corrected for multiple comparisons with the Benjamini and Hochberg method (*FDR=10%). p-values: MO p = 0.025, SS p = 0.038, AUD p = 0.082, VIS p = 0.31, ACA p = 0.050, PL p = 0.027, ILA p = 0.0086, ORB p = 0.012, GU p = 0.013, VISC p = 0.018, AI p = 0.016, RSP p = 0.20, TEA p = 0.11, PERI p = 0.074, ECT p = 0.10, OLF p = 0.017, HIP p = 0.10, RHP p = 0.052, CLA p = 0.092, EP p = 0.035, LA p = 0.028, BLA p = 0.051, BMA p = 0.036, PA p = 0.043, STRd p = 0.029, STRv p = 0.015, LSX p = 0.012, sAMY p = 0.037, PALd p = 0.036, PALv p = 0.033, PALm p = 0.024, PALc p = 0.012, DORsm p = 0.16, DORpm p = 0.095, PVZ p = 0.016, PVR p = 0.020, MEZ p = 0.029, LZ p = 0.033, MB p = 0.28, P p = 0.030, MY p = 0.054, VERM p = 0.50, HEM p = 0.043, CBN p = 0.92). Acronyms for each major brain region is listed.

Extended Data Fig. 8 Optically induced tachycardia increases *Fos* mRNA expression in brain regions associated with the central autonomic network.

a,**b**, Percentage of cells that are *Fos*$^{+}$ determined from in situ hybridization for *Fos* mRNA (magenta) following cardiac pacing in control and ChRmine-expressing mice in the nucleus of the solitary tract (NTS) (**a**) and in the locus coeruleus (LC) (**b**) (n = 4 mice per group, unpaired two-tailed t-test, ***p = 0.00029 (NTS), **p = 0.0027). We have also performed in situ hybridization for *Fos* mRNA in the nodose ganglion and observed potential induction in these vagal sensory neurons by the cardiac signals (n = 4, control 0.96% ± 0.7; ChRmine 5.29% ± 1.6, unpaired two-tailed t-test, p = 0.04). **c**, Confocal images of the locus coeruleus stained for *Fos* (magenta), the norepinephrine transporter *Slc6a2* (yellow) and DAPI (blue). **d**, Quantification of neurons that express *Slc6a2* that also co-localize with *Fos* during optical pacing (n = 4 mice per group, unpaired two-tailed t-test: ****p = 1.0e-7). Scale bar=20 µm. Data represent mean ± s.e.m.

Extended Data Fig. 9 Behavioural and physiological effects from inhibition of the posterior insula.

a, Mice were evaluated on the EPM with iC++ inhibition only during the "ON" epoch, with no optical pacing vest (n = 6 mice per group; two-way repeated-measures ANOVA: group (opsin) x time interaction $F_{(2,20)}$=2.569, p = 0.1016; group (opsin) effect $F_{(1,10)}$=0.0392, p = 0.8470; time effect $F_{(2,20)}$=0.1486, p = 0.8629). Note inhibition of pIC did not significantly alter time spent in the open arms of the EPM. **b**, Schematic overview of the pIC inhibition experiment. Mice were evaluated on the Vogel conflict task with elevated chance of shock (30%) while wearing the optical pacing vest but without receiving cardiac stimulation. Optogenetic inhibition was performed during the entire duration of the 30 min trial. **c**, Individual traces of lever presses during shock day with 30% chance of shock and bilateral pIC inhibition (n = 6 mice per group). **d**, Lever-pressing rate averaged across entire 30% shock session (n = 6 mice per group). **e**, Heart rate as a function of time during pIC inhibition (blue) (n = 3 mice per group). Data represent mean ± s.e.m.

Supplementary information

Reporting Summary

Supplementary Video 1

In vitro optical pacing of cardiomyocytes. Representative brightfield movie of contracting cardiomyocytes expressing ChRmine. Optical stimulation (light ON) was applied at 5 Hz, with 10-ms pulse width at 585 nm.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Hsueh, B., Chen, R., Jo, Y. *et al.* Cardiogenic control of affective behavioural state. *Nature* **615**, 292–299 (2023). https://doi.org/10.1038/s41586-023-05748-8

- Received: 31 December 2021
- Accepted: 20 January 2023
- Published: 01 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05748-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Anxiety can be created by the body, mouse heart study suggests

- Sara Reardon

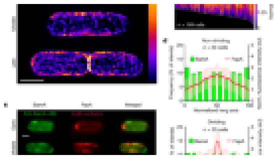
Nature News 01 Mar 2023

How an anxious heart talks to the brain

- Yoni Couderc
- Anna Beyeler

Nature News & Views 01 Mar 2023

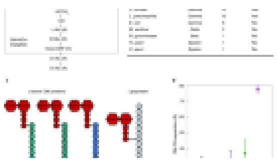
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05748-8



Peptidoglycan maturation controls outer membrane protein assembly

15 June 2022

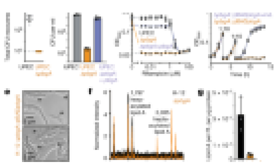
Gideon Mamou, Federico Corona, … Waldemar Vollmer



β-Barrel proteins tether the outer membrane in many Gram-negative bacteria

02 November 2020

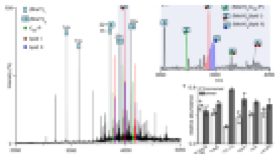
Kelsi M. Sandoz, Roger A. Moore, … Robert A. Heinzen



Structure of the essential inner membrane lipopolysaccharide–PbgA complex

12 August 2020

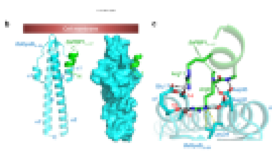
Thomas Clairfeuille, Kerry R. Buchholz, … Steven T. Rutherford



Peptidoglycan biosynthesis is driven by lipid transfer along enzyme-substrate affinity gradients

27 April 2022

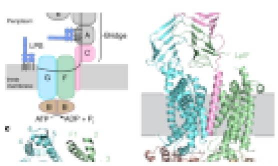
Abraham O. Oluwole, Robin A. Corey, … Carol V. Robinson



The cell cycle regulator GpsB functions as cytosolic adaptor for multiple cell wall enzymes

16 January 2019

Robert M. Cleverley, Zoe J. Rutter, … Richard J. Lewis



Structural basis of unidirectional export of lipopolysaccharide to the cell surface

20 March 2019

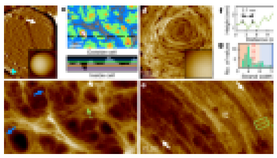
Tristan W. Owens, Rebecca J. Taylor, … Daniel Kahne



[Pushing the envelope: LPS modifications and their consequences]

29 May 2019

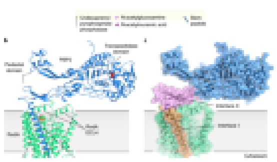
Brent W. Simpson & M. Stephen Trent



[The architecture of the Gram-positive bacterial cell wall]

29 April 2020

L. Pasquina-Lemonche, J. Burns, … J. K. Hobbs



[Structural coordination of polymerization and crosslinking by a SEDS–bPBP peptidoglycan synthase complex]

09 March 2020

Megan Sjodt, Patricia D. A. Rohs, … Andrew C. Kruse

- Article
- Open Access
- Published: 01 March 2023

Coordination of bacterial cell wall and outer membrane biosynthesis

- Katherine R. Hummels,
- Samuel P. Berry,
- Zhaoqi Li,
- Atsushi Taguchi,
- Joseph K. Min,
- Suzanne Walker,
- Debora S. Marks &
- …
- Thomas G. Bernhardt

Show authors

Nature **volume 615**, pages 300–304 (2023)

- 6868 Accesses
- 89 Altmetric
- Metrics details

Abstract

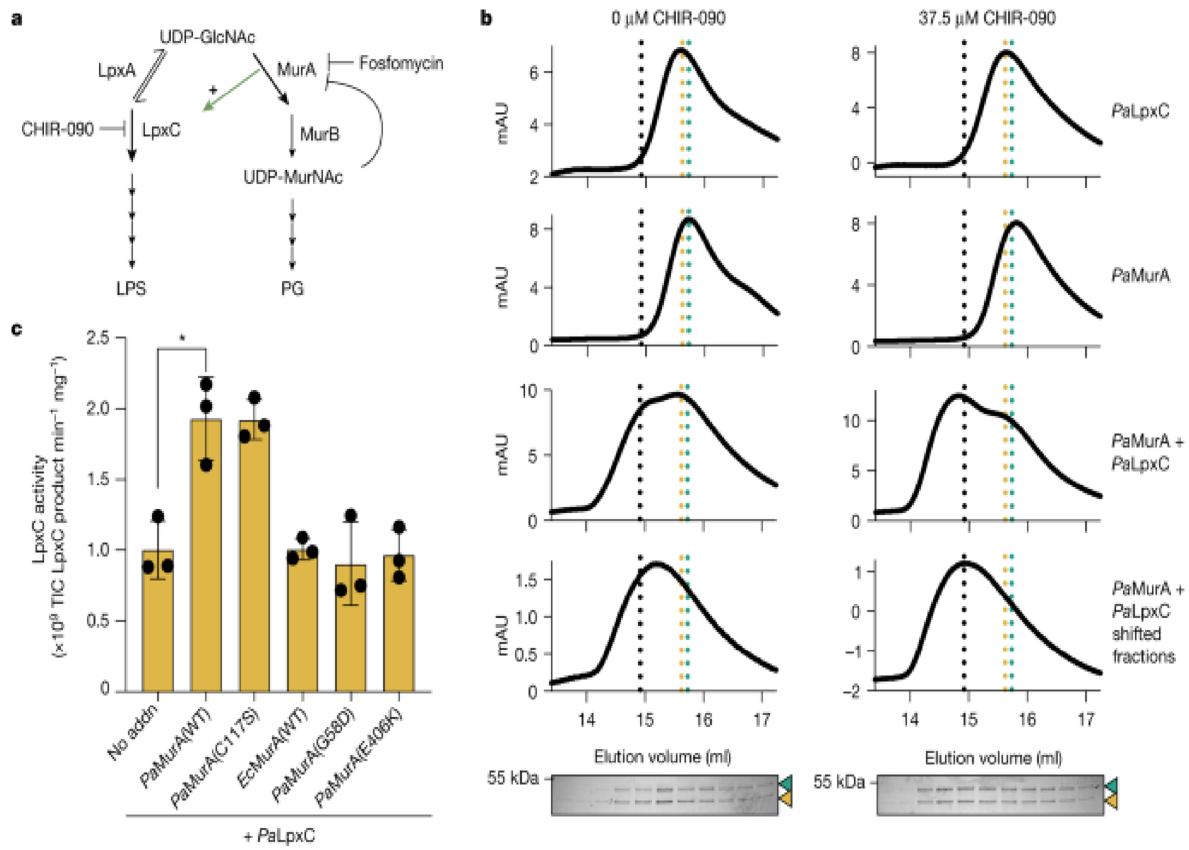
Gram-negative bacteria surround their cytoplasmic membrane with a peptidoglycan (PG) cell wall and an outer membrane (OM) with an outer leaflet composed of lipopolysaccharide (LPS)[1]. This complex envelope presents a formidable barrier to drug entry and is a major determinant of the intrinsic antibiotic resistance of these organisms[2]. The biogenesis pathways that build the surface are also targets of many of our most effective antibacterial therapies[3]. Understanding the molecular mechanisms underlying the assembly of the Gram-negative envelope therefore promises to aid the development of new treatments effective against the growing problem of drug-resistant infections. Although the individual pathways for PG and OM synthesis and assembly are well characterized, almost nothing is known about how the biogenesis of these essential surface layers is coordinated. Here we report the discovery of a regulatory interaction between the committed enzymes for the PG and LPS synthesis pathways in the Gram-negative pathogen *Pseudomonas aeruginosa*. We show that the

PG synthesis enzyme MurA interacts directly and specifically with the LPS synthesis enzyme LpxC. Moreover, MurA was shown to stimulate LpxC activity in cells and in a purified system. Our results support a model in which the assembly of the PG and OM layers in many proteobacterial species is coordinated by linking the activities of the committed enzymes in their respective synthesis pathways.

Main

The biosynthetic pathways for phospholipids, PG and LPS in Gram-negative bacteria rely on shared precursor pools (Fig. 1a). Therefore, flux through each pathway must be balanced to prevent overconsumption of essential precursors by a single pathway[4,5,6]. LPS biosynthesis requires both UDP-*N*-acetylglucosamine (UDP-GlcNAc) and acyl-ACP molecules that are also used for PG and phospholipid biosynthesis, respectively[6,7]. Additionally, overproduction of LPS results in the toxic accumulation of LPS intermediates in the inner membrane[8]. Flux through the LPS pathway must therefore be tightly regulated.

Fig. 1: *Pa*MurA interacts with and activates *Pa*LpxC.



a, Schematic representation of the biosynthetic pathways responsible for PG and LPS biosynthesis, showing the committed enzymes MurA and LpxC, and other relevant enzymes (LpxA and MurB). T-bars indicate inhibition by the antibiotics fosfomycin and CHIR-090. The green arrow indicates activation of LpxC by MurA. **b**, Size-exclusion chromatography in which 7.5 µM of purified F–LpxC and H–MurA were resolved either individually or as a mixture in the presence or absence of 37.5 µM CHIR-090 as indicated. The shifted F–LpxC + H–MurA fractions were subsequently collected, resubjected to size-exclusion chromatography, and the resulting fractions were resolved by SDS–PAGE followed by Coomassie staining. Dotted lines indicate the peak mobilities for F–LpxC (gold), H–MurA (cyan) or the shifted F–LpxC + H–MurA fractions (black) in the presence of CHIR-090. The mobilities of H–MurA and F–LpxC in SDS–PAGE assays are indicated by a cyan or gold arrowhead, respectively. Data are representative of two replicates. For gel source data, see Supplementary Fig. 1. **c**, Catalytic activity (total ion counts (TIC)) of purified *Pa*LpxC (100 nM) alone (no addition (No addn)) or in the presence of MurA variants (100 nM). Dots indicate the values obtained for three individual replicates, bars indicate the

mean, and error bars represent their standard deviation. $^*P = 0.0109$ (unpaired, two-tailed *t*-test).

Source Data

Enterobacteria such as *Escherichia coli* control LPS synthesis through regulated proteolysis of the committed enzyme, *Ec*LpxC, by FtsH[4,9]. Previous studies of LpxC in *P. aeruginosa* (*Pa*LpxC), however, showed that it was not proteolysed[10]. Accordingly, N-terminally His-tagged *Pa*LpxC (H–*Pa*LpxC) accumulated normally in an *ftsH*-deletion mutant (Extended Data Fig. 1a). Thus, LPS biogenesis in *P. aeruginosa* seems to be regulated through a mechanism distinct from that in enterobacteria.

*Pa*LpxC is activated by *Pa*MurA

The overproduction of H–*Ec*LpxC but not H–*Pa*LpxC inhibited growth of *P. aeruginosa* and increased cellular levels of LPS (Extended Data Figs. 1b and 2a), suggesting that *Pa*LpxC is regulated in *P. aeruginosa* cells through a mechanism ineffective against *Ec*LpxC. To identify possible regulatory factors, H–*Pa*LpxC or H–*Ec*LpxC was produced in *P. aeruginosa* and interaction partners were identified following affinity purification. *Pa*MurA and PA4701 were the only proteins enriched with H–*Pa*LpxC but not H–*Ec*LpxC (Supplementary Table 1). PA4701 is a non-essential protein (Extended Data Fig. 3a) of unknown function, whereas *Pa*MurA is the essential committed enzyme for PG synthesis[11,12] (Fig. 1a and Extended Data Fig. 3b,c). We focused on the *Pa*MurA–*Pa*LpxC interaction and validated it using in vivo pulldown assays with H–*Pa*LpxC and N-terminally Flag-tagged *Pa*MurA (F–*Pa*MurA; Extended Data Fig. 4a). Notably, the co-purification was enhanced when cells were treated with the LpxC inhibitor CHIR-090, suggesting that the drug-bound LpxC enzyme may have a greater affinity for MurA (Extended Data Fig. 4a). Purified His-tagged *Pa*MurA (H–*Pa*MurA; Extended Data Fig. 5a) was also specifically pulled down by Flag-tagged *Pa*LpxC (F–*Pa*LpxC) using anti-Flag resin, indicating that the interaction was direct (Extended Data Fig. 4b). Purified F–*Pa*LpxC and H–*Pa*MurA were subjected to size-exclusion chromatography (SEC) individually or as 1:1 mixtures with or without CHIR-090. In the mixed sample, a large proportion of protein eluted at a volume corresponding to the

heterodimer of F–*Pa*LpxC and H–*Pa*MurA (Fig. 1b). Consistent with the pulldown assays, re-application of the heterodimer fractions to the SEC column showed that the complex remained most stable in the presence of CHIR-090 (Fig. 1b). Notably, H–*Pa*MurA stimulated the enzymatic activity of F–*Pa*LpxC (Fig. 1c). This stimulation was specific to *Pa*MurA as the addition of H–*Ec*MurA had no effect (Fig. 1c). We conclude that *Pa*MurA is a direct and specific activator of *Pa*LpxC in vitro.

We reasoned that if *Pa*MurA is an essential activator of *Pa*LpxC in *P. aeruginosa* cells, then *Pa*MurA depletion should result in a phenotype that resembles the simultaneous inactivation of both essential enzymes. Cells treated with the MurA inhibitor fosfomycin have a PG synthesis inhibition phenotype involving membrane bleb formation and lysis[11] (Extended Data Fig. 3d–f). However, cells depleted of *Pa*MurA instead adopted an enlarged, ovoid shape (Extended Data Fig. 3e,f). This phenotype resembled that of cells treated with both CHIR-090 and fosfomycin (Extended Data Fig. 3d–f), suggesting that *Pa*MurA depletion impairs the synthesis of both PG and LPS. Accordingly, *Pa*MurA depletion reduced LPS levels, whereas depletion of the next enzyme in the pathway, *Pa*MurB, did not (Extended Data Figs. 2b and 3b,c) and instead caused a terminal phenotype resembling that following fosfomycin treatment (Extended Data Fig. 3d–f). We conclude that *Pa*MurA is required for the biosynthesis of normal cellular levels of LPS.

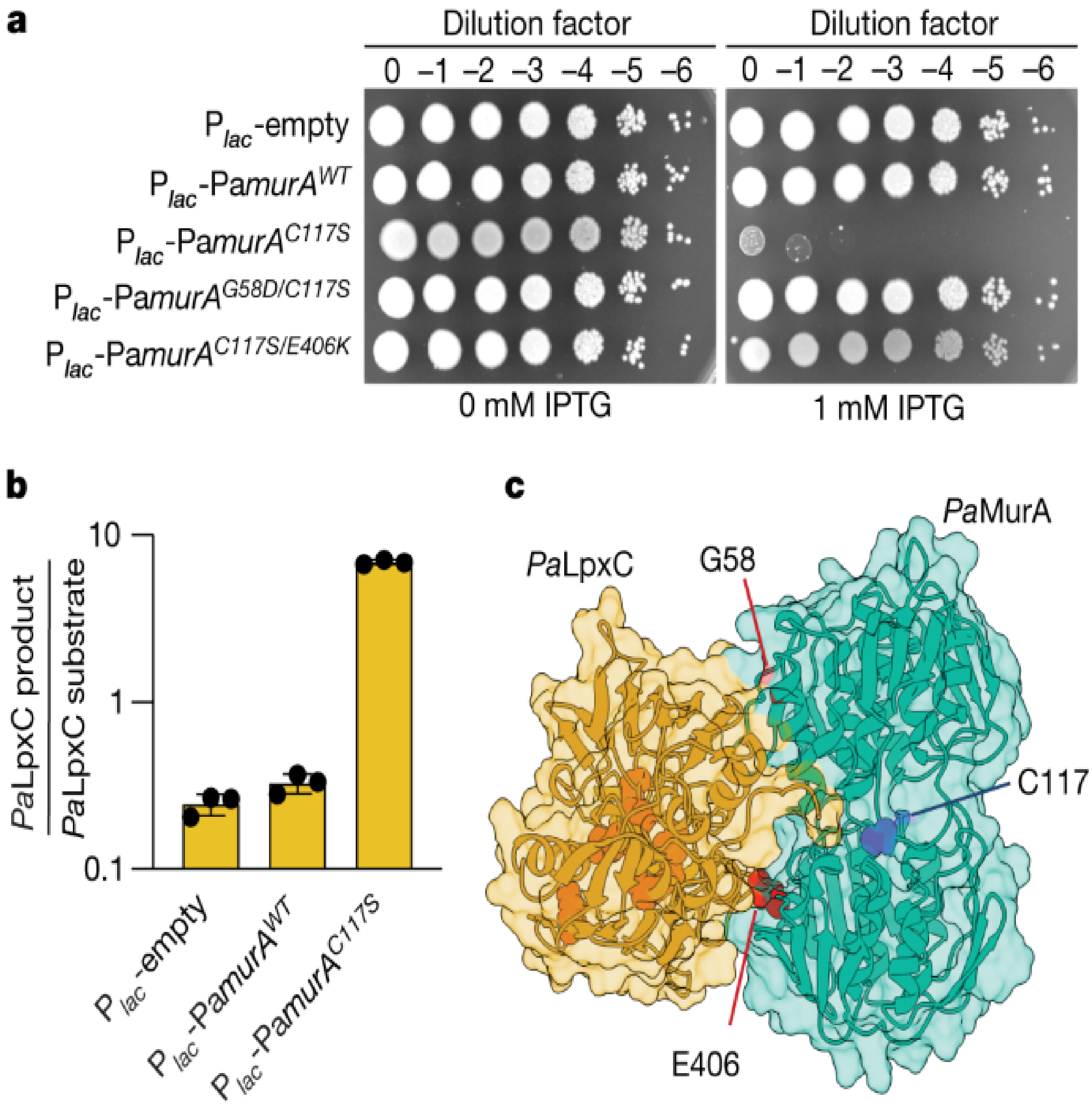
***Pa*MurA has two essential functions**

Overproduction of wild-type (WT) *Pa*MurA did not increase LPS levels as expected for an activator of *Pa*LpxC (Extended Data Fig. 2c). We suspected this result might be due to the enzymatic activity of *Pa*MurA competing with the LPS pathway for UDP-GlcNAc (Fig. 1a) thereby preventing runaway LPS synthesis. Alternatively, a particular MurA conformer[13] may be the activator and it may not be produced in sufficient levels to stimulate LPS synthesis on *Pa*MurA overexpression. These scenarios predict that the overproduction of catalytically inactive or conformationally trapped *Pa*MurA variants should cause lethal levels of LPS production. We therefore searched for toxic *Pa*MurA variants. *P. aeruginosa* was transformed with a plasmid encoding mutagenized Pa*murA* under the control of an isopropyl-β-

d-thiogalactoside (IPTG)-regulated promoter. The resulting library was pooled and grown in liquid medium with inducer. Plasmids that caused lysis were purified from the culture supernatant, and the Pa*murA* genes from those causing an IPTG-dependent growth defect on retransformation were sequenced.

Twenty-three toxic *Pa*MurA variants (designated *Pa*MurA*) were identified (Fig. 2a and Extended Data Fig. 6a). Their growth inhibitory activity was alleviated by a normally lethal concentration of CHIR-090, suggesting that they hyperactivate *Pa*LpxC (Extended Data Fig. 6a). The amino acid changes in the *Pa*MurA* variants mapped around the active site of the enzyme[14] (Extended Data Fig. 6b) and included the catalytic cysteine residue (C117) that forms a covalent intermediate with the phosphoenolpyruvate substrate[15]. A subset of *Pa*MurA* variants were purified (Extended Data Fig. 5a) and found to have markedly reduced enzymatic activity while retaining their ability to activate purified *Pa*LpxC (Extended Data Fig. 6c,d).

Fig. 2: *Pa*MurA(C117S) is a potent activator of *Pa*LpxC in vivo.



a, Viability assay in which serial dilutions of PAO1 harbouring an empty plasmid or the indicated Pa*murA* variant under IPTG-inducible control were plated on LB agar with or without IPTG supplementation as indicated. Data are representative of three biological replicates. **b**, Ratio of LpxC product to substrate detected by liquid chromatography–mass spectrometry in the aqueous fraction of methanol–chloroform-extracted *P. aeruginosa* whole-cell lysates. Dots indicate the values obtained for three biological replicates, bars indicate the mean, and error bars represent their standard deviation. **c**, Model structure of the *Pa*LpxC–*Pa*MurA complex predicted by AlphaFold[19]. *Pa*LpxC is represented in gold and *Pa*MurA is represented in

cyan. *Pa*MurA residues G58 and E406 are highlighted in red and residue C117 is highlighted in navy. *Pa*LpxC active-site residues[34] are depicted in orange.

Source Data

The equivalent of the C117S substitution in *Pa*MurA has been well characterized in other orthologues in which it has been shown to trap the enzyme in a closed conformation bound to its product[16,17]. We therefore chose to study the potential in vivo activation of *Pa*LpxC by *Pa*MurA(C117S) further. The LpxC substrate (UDP-3-*O*-(*R*-3-hydroxydecanoyl)-*N*-acetylglucosamine) and product (UDP-3-*O*-(*R*-3-hydroxydecanoyl)-glucosamine) were extracted and quantified from cells overproducing *Pa*MurA(WT) or *Pa*MurA(C117S). The *Pa*LpxC product/substrate ratio was unchanged in cells overproducing *Pa*MurA(WT) relative to an empty vector control (Fig. 2b and Extended Data Fig. 7). However, overproduction of *Pa*MurA(C117S) led to a 100-fold increase in the *Pa*LpxC product/substrate ratio and increased LPS levels, indicating that this variant is a potent activator of *Pa*LpxC in vivo (Fig. 2b and Extended Data Figs. 2c and 7). We therefore conclude that *Pa*MurA activates *Pa*LpxC in cells but that only catalytically inactive *Pa*MurA variants promote toxic levels of LPS synthesis when they are overproduced owing to their inability to compete with the LPS synthesis pathway for UDP-GlcNAc.

To further characterize the interaction between *Pa*LpxC and *Pa*MurA, we searched for non-toxic derivatives of *Pa*MurA(C117S), reasoning that some would be unable to bind *Pa*LpxC. Survivors of F–*Pa*MurA(C117S) production from a mutagenized plasmid were selected followed by an immunoblot-based screen for isolates producing stable, full-length F–*Pa*MurA(C117S). This procedure identified *Pa*MurA(C117S/G58D) and *Pa*MurA(C117S/E406K). We confirmed that both variants were not toxic when overproduced to levels similar to *Pa*MurA(C117S) (Fig. 2a and Extended Data Fig. 8a). We next purified H–*Pa*MurA(G58D) and H–*Pa*MurA(E406K) variants with intact active sites (Extended Data Fig. 5a). Both retained MurA activity (Extended Data Fig. 8b), and consistent with the genetic results, they failed to activate F–*Pa*LpxC in vitro (Fig. 1c). Notably, however, whereas H–*Pa*MurA(G58D) was unable to bind to F–*Pa*LpxC, H–*Pa*MurA(E406K) exhibited binding activity similar to that of

H–*Pa*MurA(WT) (Extended Data Fig. 8c). Thus, the G58D change impairs the ability of *Pa*MurA to bind *Pa*LpxC, whereas the E406K substitution seems to disrupt the activation mechanism following complex formation. The residues G58 and E406 both lie on the same surface of the *Pa*MurA structure, suggesting that this face comprises the *Pa*LpxC-binding interface (Fig. 2c).

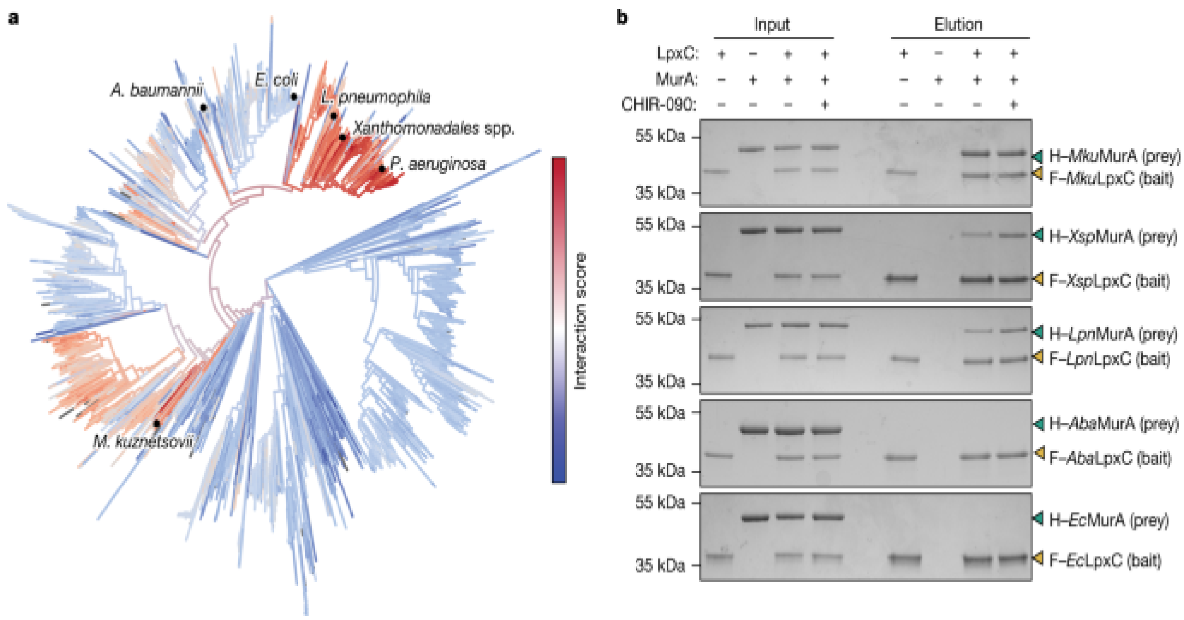
The results thus far suggest that *Pa*MurA may have two essential functions: PG synthesis and activation of *Pa*LpxC. This model predicts that a catalytically active variant of *Pa*MurA that cannot activate *Pa*LpxC should fail to complement a Pa*murA* deletion but remain capable of complementing a catalytically dead *Pa*MurA variant that retains its *Pa*LpxC activation function. To test this prediction, Pa$murA^{WT}$ was placed under P_{lac} control and the native allele was either deleted or converted to Pa$murA^{C117S}$. We then introduced a plasmid with or without the Pa$murA^{G58D}$ gene controlled by an arabinose-inducible promoter (P_{ara}). As expected, *Pa*MurA(WT) depletion resulted in a severe growth defect in either the Δ*murA* or Pa$murA^{C117S}$ backgrounds (Extended Data Fig. 8d). Notably, however, the production of *Pa*MurA(G58D) was able to restore growth on *Pa*MurA(WT) depletion in the context of the Pa$murA^{C117S}$ allele, but not the *murA* deletion (Extended Data Fig. 8d). We therefore conclude that the *Pa*LpxC activation function of *Pa*MurA is essential and separable from its catalytic activity.

LpxC–MurA interaction is conserved

To investigate the conservation of the LpxC–MurA regulatory interaction, we used evolutionary covariation analysis[18]. However, because both enzymes are conserved throughout Gram-negative bacteria but only a subset are likely to interact, we could not reliably detect residues that covary between LpxC and MurA without first knowing which regions of the two proteins interact; non-interacting pairs among the genomes analysed generated too much ‘noise’ to detect a clear interaction signature. We therefore modelled the structure of the LpxC–MurA complex with AlphaFold[19], which predicted a high-confidence structure for *Pa*LpxC–*Pa*MurA but not between the corresponding *E. coli* proteins (Fig. 2c and Extended Data Fig. 9a,b). The G58 and E406 residues of *Pa*MurA

implicated in *Pa*LpxC binding and/or activation lie at the modelled interaction interface (Fig. 2c), supporting the accuracy of the structure prediction. Using this model as a guide, evolutionary covariation analysis of residues predicted to be at the interaction interface was carried out on LpxC–MurA pairs from 8,302 proteobacterial genomes. Each LpxC–MurA pair was then assigned an ‘interaction score’ as an indication of how strongly the two proteins were predicted to interact (Fig. 3a).

Fig. 3: Direct interaction between LpxC and MurA is observed in diverse bacteria.



a, MurA–LpxC interaction scores calculated from a direct coupling analysis-based approach (Methods) plotted on a maximum-likelihood phylogenetic tree based on concatenated MurA and LpxC sequences generated with FastTree. Experimentally tested interactions are highlighted. **b**, Purified F–LpxC (2.5 μM) was mixed in a 1:1 ratio with purified H–MurA in the presence or absence of CHIR-090 (5.7 μM) as indicated. The mixtures were pulled down with anti-Flag resin and the input and eluate were subjected to SDS–PAGE and Coomassie staining. Mobilities of H–MurA and F–LpxC are indicated by a cyan or gold arrowhead, respectively. Data are representative of at least two replicates. For gel source data, see Supplementary Fig. 1.

Source Data

The *P. aeruginosa* LpxC–MurA pair had a high interaction score as expected, as did those from other Pseudomonadales and a subset of other gammaproteobacteria including the Legionalles, Xanthomonadales and Oceanospirillales orders (Fig. 3a). In addition, high interaction scores were also observed in a subset of Rhodospirillales, an order of alphaproteobacteria that is highly divergent from *P. aeruginosa* (Fig. 3a). On the basis of the distribution of interaction scores, we estimate that 48% of gammaproteobacteria and 35% of alphaproteobacteria encode LpxC–MurA pairs that are capable of interacting (Extended Data Fig. 9e). To investigate the accuracy of the covariation analysis, three additional LpxC–MurA pairs with high interaction scores (*Magnetospirillum kuznetsovii*, *Xanthomonadales* spp. and *Legionella pneumophila*) and two LpxC–MurA pairs with low interaction scores (*E. coli* and *Acinetobacter baumannii*) were purified (Extended Data Fig. 5a) and their ability to interact was tested in vitro. Only those pairs with high interaction scores were observed to interact (Fig. 3b). Furthermore, *L. pneumophila* LpxC (*Lpn*LpxC), but not *Ec*LpxC, exhibited increased activity in vitro in the presence of its cognate MurA (Extended Data Fig. 9f,g). Thus, we conclude that LpxC and MurA from a variety of proteobacteria interact, suggesting that the regulatory link between the PG and LPS biogenesis pathways is conserved in a variety of Gram-negative bacteria.

Model for balanced LPS and PG synthesis

The LPS biosynthetic pathway shares critical precursors with both the phospholipid and PG biosynthetic pathways (Fig. 1a). In *E. coli* and other enterobacteria, the need for balanced consumption of acyl-ACP molecules by the LPS and phospholipid biosynthetic pathways has been well documented[4,20,21,22,23]. By contrast, little is known about how UDP-GlcNAc utilization is coordinated by the LPS and PG biosynthetic pathways to allow uniform cell envelope expansion. We propose that the *Pa*LpxC-activating function of *Pa*MurA serves to limit LPS biosynthesis such that it cannot outrun PG biosynthesis. In support of this model, overexpression of *Pa*MurA(WT), which is capable of competing with the LPS synthesis pathway for UDP-GlcNAc, had no impact on cellular viability (Fig. 2a). Overexpression of the catalytically inactive variant *Pa*MurA(C117S), however, resulted in the imbalanced activation of *Pa*LpxC in vivo, resulting

in cell death (Fig. 2a). By limiting the catalytically active population of *Pa*LpxC to *Pa*MurA levels, runaway LPS biosynthesis is not possible. On the other hand, MurA has been shown to be feedback inhibited by the downstream PG precursor UDP-MurNAc[24]. Under conditions in which the flux of PG precursor synthesis is too high, UDP-MurNAc will accumulate such that it binds to MurA and locks it into a closed conformation similar to that of *Pa*MurA(C117S)[13,16]. Thus, when PG biosynthesis outpaces LPS biosynthesis, MurA will be subject to feedback inhibition, decreasing its catalytic activity without affecting its ability to activate *Pa*LpxC to stimulate LPS biosynthesis and rebalance the pathways.

In enterobacteria, LpxC is proteolysed by FtsH, which is in turn regulated by at least two accessory proteins: LapB and YejM[5,8,25,26]. Although FtsH is broadly conserved, LapB and YejM are observed only in a subset of proteobacterial genomes. By contrast, LpxC and MurA are nearly universally conserved in proteobacteria. Our computational analysis suggests that a substantial group of gammaproteobacteria and alphaproteobacteria control LpxC through activation by MurA (Fig. 3a). Why bacteria use different strategies to regulate flux through the LPS biosynthetic pathway is unclear. We suggest, however, that it may reflect differences in their environmental niche. For example, *P. aeruginosa* preferentially utilizes tricarboxylic acids over carbohydrates as a carbon source[27]. Thus, activated sugars such as UDP-GlcNAc may be more limiting for *P. aeruginosa* than for sugar-loving *E. coli* such that it focuses its regulation of LpxC on the appropriate partitioning of UDP-GlcNAc. It would not be surprising if future studies uncover other as-yet-unknown strategies of LpxC regulation in other organisms.

Conclusions

In conclusion, our work has identified a previously unknown and unexpected regulatory interaction between essential enzymes involved in the production of two different layers of the Gram-negative cell envelope. In addition to uncovering a potential mechanism for coordinating the biogenesis of the cell wall and OM, these findings also have implications for antibiotic development. MurA is the target of the antibiotic fosfomycin and LpxC has been the subject of several drug discovery campaigns that have

yielded inhibitors with potent antibacterial activity[11,28,29,30,31,32,33]. The connection between these enzymes identified here suggests that it may be possible to develop dual-targeting drugs that alter both MurA and LpxC activity to simultaneously disrupt PG and OM assembly to kill *P. aeruginosa* and/or sensitize it to other antibiotics made ineffective by the barrier function of its envelope. These findings also raise the possibility that there are many more regulatory interactions between enzymes involved in the biogenesis of different cell envelope components waiting to be discovered.

Methods

Media, bacterial strains and plasmids

As indicated, cells were grown in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl) or minimal M9 medium[35] supplemented with 0.2% casamino acids and 0.2% glucose. The following concentrations of antibiotics were used: chloramphenicol, 25 μg ml^{-1}; kanamycin, 25 μg ml^{-1}; gentamycin, 30 μg ml^{-1} (*P. aeruginosa*) or 15 μg ml^{-1} (*E. coli*); carbenicillin, 200 μg ml^{-1} (*P. aeruginosa*) or 50 μg ml^{-1} (*E. coli*). Plasmids used in this study are listed in Supplementary Table 2. The bacterial strains used in this study are listed in Supplementary Table 3. All *P. aeruginosa* strains used in the reported experiments are derivatives of PAO1. Primers used in this study are listed in Supplementary Table 4. Unless stated otherwise, polymerase chain reaction (PCR) was carried out using Q5 polymerase (NEB M0492L) for cloning purposes and GoTaq Green DNA polymerase (Promega M7123) for diagnostic purposes, both according to the manufacturer's instructions. Plasmid DNA and PCR fragments were purified using the PurePlasmid miniprep kit (CW Biosciences CW0500M) or the DNA Clean-up kit (CW Biosciences CW2301M), respectively. Details on strain and plasmid construction can be found in the Supplementary Information. No statistical methods were used to predetermine sample sizes used for experiments, but sample sizes are in line with field standards.

Isolation of toxic *murA* alleles

Alleles of *murA* that are toxic when overexpressed were isolated as described previously[36] with slight modifications. First, the *murA* gene was mutagenized by PCR with Taq polymerase (NEB M0267L) from purified pKH37 plasmid using primer pair 15/76 in five separate reactions. A 500 ng quantity of each of the five resulting *murA** pools was then separately used as a megaprimer to amplify 50 ng of pKH37 using Q5 polymerase (NEB M0492L) in 50-µl reactions with the following thermocycler settings: 1) 95 °C for 3 min; 2) 95 °C for 50 s; 3) 60 °C for 50 s; 4) 72 °C for 10 min; 5) repeat steps 2–4 for a total of 25 cycles. Twenty units of DpnI (NEB R0176L) was subsequently added to each reaction and digestion of unamplified DNA was allowed to proceed at 37 °C for 1.5 h. Each reaction was drop dialysed using 0.025-µm mixed cellulose ester membranes (Millipore VSWP02500) floated on Milli-Q water for 20 min. A 12 µl volume of each dialysed product was then separately transformed into 100 µl of NEB-5-alpha electrocompetent *E. coli* (NEB C2989K), and recovered in SOC outgrowth medium (NEB B9020S), and transformants were selected by plating the outgrowth onto LB agar supplemented with gentamycin. Approximately 1 million colonies from each of the five libraries were separately pooled and the pKH37 derivatives were purified using the PurePlasmid miniprep kit (CW Biosciences CW0500M).

PAO1 was grown in 25 ml LB overnight at 37 °C, centrifuged at 12,000*g* for 5 min, and resuspended in an equal volume of 300 mM sucrose. The centrifugation and resuspension steps were repeated for a total of four washes. After a final centrifugation at 12,000*g* for 5 min, the cell pellet was resuspended in 1/20 of the original volume. A 1.2 µg quantity of pKH37-derived libraries was separately transformed into 150 µl of electrocompetent PAO1. Transformation reactions were recovered in LB at 37 °C for 1 h and transformants were selected by plating the outgrowth onto LB agar supplemented with 30 µg ml^{-1} gentamycin. Approximately 1–2 million colonies from each library were separately pooled.

To identify pKH37 variants that harbour a toxic allele of *murA*, each PAO1 × pKH37* library was diluted to an optical density at 600 nm (OD_{600nm}) of 0.01 in 3 ml LB supplemented with 30 µg ml^{-1} gentamycin and grown at 37 °C for 2 h 45 min. IPTG was added to each culture to a final concentration of 1 mM and the incubation at 37 °C was continued for an additional 2 h. Each

culture was then subjected to centrifugation at 21,130*g* for 5 min and released DNA from each library was separately purified from 2.5 ml of the supernatant using the DNA clean-up kit (CW Biosciences CW2301M). The purified DNA was subsequently re-transformed into PAO1 electrocompetent cells prepared as described above and transformants were selected by plating the outgrowth on LB agar supplemented with gentamycin. The resulting colonies were then patched onto LB agar plates supplemented with gentamycin with and without 1 mM IPTG and grown overnight at 37 °C. The *murA* allele encoded by IPTG-sensitive isolates was determined by Sanger sequencing using primer pair 15/76.

Isolation of MurA(C117S) loss-of-toxicity variants

pKH100 was mutagenized by passaging it through *E. coli* strain XL1-red. pKH100 was transformed into XL1-red and five separate cultures were inoculated, each using four unique transformants. After propagation overnight, plasmids were purified using the PurePlasmid miniprep kit (CW Biosciences CW0500M) to create five pKH100* libraries. PAO1 was grown in 25 ml LB overnight at 37 °C, centrifuged at 12,000*g* for 5 min, and resuspended in an equal volume of 300 mM sucrose. The centrifugation and resuspension steps were repeated for a total of four washes. After a final centrifugation at 12,000*g* for 5 min, the cell pellet was resuspended in 1/20 of the original volume. A 1.2 µg quantity of each pKH100* was separately transformed into 150 µl of electrocompetent PAO1. Transformation reactions were recovered in LB at 37 °C for 1 h and transformants were selected by plating the outgrowth onto LB agar supplemented with gentamycin. Approximately 1–10 million colonies from each library were separately pooled and stored in LB + 10% dimethylsulfoxide (DMSO) at −80 °C.

PAO1 × pKH100* libraries were plated on LB agar supplemented with 30 µg ml^{-1} gentamycin and 1 mM IPTG, which allowed growth of only approximately 0.3–1% of all colony-forming units. A total of 17 or 18 IPTG-resistant colonies from each pool were streak-purified and the expression level and molecular weight of F–MurA(C117S)* from each isolate were determined by western blot analysis after growth in LB and induction for 1 h in mid-log phase. Only those isolates with expression

levels and molecular weights consistent with F–MurA(C117S) were subjected to Sanger sequencing with primers 15 and 76. In three of the five libraries analysed, *F–murA*^{WT} was recovered. *F–murA*$^{C117S/G58D}$ and *F–murA*$^{C117S/E406K}$ were each recovered in one of the five libraries.

Protein purification

Details on protein purification can be found in the Supplementary Information.

Viability assays

Overnight cultures grown at 30 °C or 37 °C were centrifuged at 12,000*g* for 2 min and resuspended to an OD_{600nm} of 1 in LB. Resuspensions were serially diluted in LB to 10^{-6} and 5 µl of each dilution was spotted onto LB agar supplemented with the appropriate inducers and antibiotics. Plates were incubated at 37 °C overnight and imaged using a Nikon D3400 camera equipped with a Nikon AF-S micro NIKKOR 40-mm lens.

Western blot analysis

Overnight cultures were diluted to an OD_{600nm} of 0.01 in 3 ml LB, grown at 37 °C for 4 h, and 2 ml was subjected to centrifugation at 12,000*g* for 3 min. When necessary, IPTG was added to 1 mM or arabinose was added to 0.1% 1 h before collecting the cultures. The cell pellet was resuspended to an OD_{600nm} of 20 in 1× SDS sample buffer (50 mM Tris pH 6.8, 10% glycerol, 2% SDS, 0.2% bromophenol blue, 1% β-mercaptoethanol) and boiled at 100 °C for 10 min. The lysate was subsequently sonicated using a Qsonica Q800R3 sonicator with 30 cycles of 1 s on, 1 s off at 25% amplitude. Protein concentration was quantified using the Non-Interfering protein assay kit (G-biosciences 786-005). A 10 µg quantity of protein was loaded onto a 10% polyacrylamide–SDS gel and subjected to electrophoresis at 150 V for 70 min. Protein was transferred to PVDF membranes using the mixed molecular weight protocol on a Bio-Rad Trans-Blot Turbo transfer system and membranes were blocked in TBST (20 mM Tris pH 8.0, 200 mM NaCl, 0.05% Tween-20) with 5% milk at room temperature for 1 h. Membranes

were then probed with anti-His (GenScript A00186-100), anti-Flag (Sigma F7425) or anti-RpoA (BioLegend 663104) diluted 1:3,000, 1:3,000 or 1:100,000, respectively, in TBST + 5% milk at room temperature for 1 h. Membranes were washed three times for 5 min each in TBST and probed with peroxidase-conjugated goat-anti-mouse (Rockland 610-1302) or rabbit TrueBlot: anti-rabbit IgG HRP (Rockland 18-8816-33) diluted 1:3,000 in TBST + 5% milk at room temperature for 1 h. Membranes were washed three times for 5 min each in TBST, developed with SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo 34580), and imaged using an Azure Biosystems C600 imager.

In vivo co-affinity purification

To identify in vivo interacting partners of *Pa*LpxC in an unbiased manner, two co-affinity purification schemes were used in replicate samples. Specifically, in condition 1, 150 mM NaCl was used in lysis and wash buffers, but in condition 2, 300 mM NaCl was used in lysis and wash buffers. Overnight *P. aeruginosa* cultures were diluted to OD_{600nm} = 0.01 in 500 ml LB and grown at 37 °C to early logarithmic phase. In the case of PA1009, the culture was induced with 0.5% arabinose 1 h before collection. The entire culture was subjected to centrifugation at 13,000*g* for 10 min and the pellets were resuspended in 4 ml of lysis buffer (50 mM Tris pH 7.5, 2% glycerol, 5 mM imidazole, 1% Triton X-100, and 150 mM or 300 mM NaCl) and cells were lysed by sonication with a Q125 Qsonica sonicator. Cell debris was pelleted by centrifugation at 21,130*g* for 15 min at 4 °C and the clarified supernatant was mixed with Ni-NTA resin (Qiagen 30230) rotating end-over-end at 4 °C for 1.5 h. Resin was washed four times with 6 ml of wash buffer (50 mM Tris pH 7.5, 2% glycerol, 25 mM imidazole, 1% Triton X-100, and 150 mM or 300 mM NaCl). Proteins were eluted from the resin in elution buffer (50 mM Tris pH 7.5, 2% glycerol, 250 mM imidazole, 0.1% Triton X-100 and 150 mM NaCl). Eluates were mixed 1:1 with 2× SDS sample buffer (100 mM Tris pH 6.8, 20% glycerol, 4% SDS, 0.4% bromophenol blue), and 40 µl was loaded onto 4–20% polyacrylamide Mini-PROTEAN TGX gels (Bio-Rad 4568094) and subjected to electrophoresis at 80 V for 25 min. The gel was subsequently stained with Coomassie R-250 and entire lanes were excised for mass spectrometry analysis.

Excised lanes were cut into approximately 1-mm^3 pieces. Gel pieces were then subjected to a modified in-gel trypsin digestion procedure[37]. Gel pieces were washed and dehydrated with acetonitrile for 10 min, followed by removal of acetonitrile. Pieces were then completely dried in a speed-vac. Gel pieces were rehydrated with 50 mM ammonium bicarbonate solution containing 12.5 ng μl^{-1} modified sequencing-grade trypsin (Promega) at 4 °C. After 45 min, the excess trypsin solution was removed and replaced with 50 mM ammonium bicarbonate solution to just cover the gel pieces. Samples were then placed in a 37 °C room overnight. Peptides were later extracted by removing the ammonium bicarbonate solution, followed by one wash with a solution containing 50% acetonitrile and 1% formic acid. The extracts were then dried in a speed-vac (about 1 h). The samples were then stored at 4 °C until analysis.

On the day of analysis, the samples were reconstituted in 5–10 μl of high-performance liquid chromatography (HPLC) solvent A (2.5% acetonitrile, 0.1% formic acid). A nanoscale reversed-phase HPLC capillary column was created by packing 2.6-μm C18 spherical silica beads into a fused silica capillary (100 μm inner diameter × about 30 cm length) with a flame-drawn tip[38]. After equilibrating the column, each sample was loaded using a Famos auto sampler (LC Packings) onto the column. A gradient was formed, and peptides were eluted with increasing concentrations of solvent B (97.5% acetonitrile, 0.1% formic acid).

As peptides eluted, they were subjected to electrospray ionization and then entered into an LTQ Orbitrap Velos Pro ion-trap mass spectrometer (Thermo Fisher Scientific). Peptides were detected, isolated and fragmented to produce a tandem mass spectrum of specific fragment ions for each peptide. Peptide sequences (and hence protein identity) were determined by matching protein databases with the acquired fragmentation pattern by the software program Sequest v28 (rev. 13; Thermo Fisher Scientific)[39]. All databases include a reversed version of all the sequences, and the data were filtered to a peptide false-discovery rate of 1–2%. Only proteins that exhibited the following criteria were reported: at least five unique peptides detected in both PA1018 samples; and at least threefold enrichment in both PA1018 samples compared to the corresponding PAO1 and PA1009 samples.

In vivo targeted pulldowns

Overnight cultures of *P. aeruginosa* strains encoding combinations of H–*Pa*LpxC, F–*Pa*MurA(WT) and F–*Pa*MurA(C117S) were diluted to OD_{600nm} = 0.01 in 50 ml LB and grown at 37 °C for 2.5 h. IPTG was added to each culture to a final concentration of 1 mM and cultures were grown for an additional 30 min at 37 °C before the addition of DMSO to 0.1% and CHIR-090 to 0.5 μg ml^{-1} where indicated. Cultures were incubated at 37 °C for an additional hour and collected by centrifugation at 12,000*g* for 10 min. Cells were washed in 1 ml LB and pelleted by centrifugation at 12,000*g* for 2 min. Cells were resuspended in 1 ml lysis buffer (25 mM Tris pH 8.0, 150 mM NaCl, 2% glycerol, 0.1% Triton X-100, 5 mM imidazole, 0.1% DMSO with or without 0.5 μg ml^{-1} CHIR-090 as indicated) and lysed by sonication on ice at 40% amplitude with four cycles of 15 s on, 15 s off. The extracts were clarified by two consecutive centrifugation steps at 21,130*g* for 5 min at 4 °C and protein concentration was determined with the Non-Interfering protein assay kit (G-biosciences 786-005). Supernatants were diluted to 3.5 mg ml^{-1} protein in 500 μl lysis buffer and mixed with 25 μl packed Ni-NTA resin (Qiagen 30230). Mixtures were incubated at 4 °C for 3 h rotating end-over-end. Beads were washed three times with 500 μl wash buffer (25 mM Tris pH 8.0, 150 mM NaCl, 2% glycerol, 0.1% Triton X-100, 25 mM imidazole, 0.1% DMSO with or without 0.5 μg ml^{-1} CHIR-090 as indicated). Protein was eluted in 40 μl elution buffer (25 mM Tris pH 8.0, 150 mM NaCl, 2% glycerol, 0.1% Triton X-100, 500 mM imidazole).

A 10 μg quantity of protein per input sample and 10 μl of eluate samples were resolved by SDS–PAGE in 10% polyacrylamide gels. Protein was transferred to PVDF membranes using the mixed molecular weight protocol on a Bio-Rad Trans-Blot Turbo transfer system and membranes were blocked in TBST + 5% milk at room temperature for 1 h. Membranes were then probed with anti-His (GenScript A00186-100) or anti-Flag (Sigma-Aldrich F7425) diluted 1:3,000 in TBST + 5% milk at room temperature for 1 h. Membranes were washed three times for 5 min each in TBST and probed with peroxidase-conjugated goat-anti-mouse (Rockland 610-1302) or rabbit TrueBlot: anti-rabbit IgG HRP (Rockland 18-8816-33) diluted 1:3,000 in TBST + 5% milk at room temperature for 1 h. Membranes were washed three times for 5 min each in TBST, developed with SuperSignal West Pico

PLUS Chemiluminescent Substrate (Thermo 34580), and imaged using an Azure Biosystems C600 imager.

In vitro co-affinity purification

Purified F–*Pa*LpxC and H–MurA variants were each diluted to 2.5 µM in co-purification buffer (25 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 1 mM dithiothreitol (DTT), 0.5% DMSO) to a final volume of 100 µl. When indicated, CHIR-090 was added to 5.7 µM. Mixtures were incubated on ice for 30 min after which 85 µl of each mixture was added to an equal volume of Anti-Flag M2 Magnetic Beads (Sigma-Aldrich M8823) and incubated at 4 °C for 1 h rotating end-over-end. Beads were collected with a magnetic rack and washed twice with 300 µl co-purification buffer and once with 200 µl co-purification buffer. When indicated, 5.7 µM CHIR-090 was maintained in the wash buffers. Proteins were eluted with 85 µl elution buffer (25 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 1 mM DTT, 100 µg µl^{-1} Flag peptide (Millipore Sigma F3290)) by incubation at room temperature for 30 min rotating end-over-end. Input samples (5 µl) and eluate samples (15 µl) were resolved by SDS–PAGE in 4–20% polyacrylamide Mini-PROTEAN TGX gels (Bio-Rad 4561095) and protein was detected with Coomassie staining.

SEC

F–*Pa*LpxC and H–*Pa*MurA(WT) either individually or in combination were diluted to 7.5 µM each in 500 µl of SEC buffer 1 (10% glycerol, 0.33% DMSO, 1 mM DTT, 25 mM Tris pH 8.0 and 150 mM NaCl). When indicated, CHIR-090 was added to 37.5 µM. The mixtures were subsequently subjected to SEC using an ActaPure system equipped with a Superdex 200 10/300 GL column equilibrated with SEC buffer 2 (10% glycerol, 1 mM DTT, 25 mM Tris pH 8.0 and 150 mM NaCl) and eluate fractions were collected every 0.35 ml. For F–*Pa*LpxC and H–*Pa*MurA(WT) mixtures, three fractions corresponding to the shifted peak (14.1–15.2 ml with no CHIR-090, 13.8–14.8 ml for with CHIR-090) were pooled and concentrated by centrifugation at 4 °C using centrifugal filters with a molecular weight cutoff of 10-kDa (Amicon UFC801024). The concentrated protein was resubjected to SEC on a Superdex 200 10/300 GL column as

described above. A 15 µl volume of the relevant fractions was resolved on a 4–20% polyacrylamide Mini-PROTEAN TGX gels (Bio-Rad 4568094) and protein was detected with Coomassie staining.

Growth curves

Cultures grown overnight at 30 °C were pelleted at 12,000*g* for 2 min and resuspended in fresh LB. Cultures were diluted in 96-well microtitre plates to an OD_{600nm} of 0.01 in 200 µl LB. Cultures were grown in a Tecan Infinite M-Plex plate reader at 37 °C and OD_{600nm} was monitored every 4 min with shaking at a 1-mm orbital for 200 s after each time point.

LPS silver stain

Overnight cultures were diluted to an OD_{600nm} of 0.01 in 25 ml LB cultures and grown for 4 h at 37 °C. For MurA- and MurB-depletion strains grown in the presence of inducer, 1 mM IPTG was present during the entire outgrowth. For MurA and LpxC overexpression strains, 1 mM IPTG or 0.1% arabinose was added after 3 h of outgrowth and incubation was continued for one additional hour. A 20 ml volume of each culture was centrifuged at 12,000*g* for 5 min and the pellet was washed in 1 ml of LB before centrifugation at 12,000*g* for 2 min. Pellets were resuspended to an OD_{600nm} of 20 in LDS sample buffer (Invitrogen NP0008) + 4% β-mercaptoethanol and boiled at 100 °C for 10 min. Samples were sonicated with a Q125 Qsonica sonicator at 25% amplitude for 1 s on, 1 s off for 30 cycles and protein concentration was determined using the Non-Interfering protein assay kit (G-biosciences 786-005). A 5 µg quantity of protein was loaded onto a 15% polyacrylamide–SDS gel and subjected to electrophoresis at 150 V for 70 min before western blot analysis of RpoA as described above. For the LPS gel, 50 µl of each sample was mixed with 1.25 µl of proteinase K (NEB P8107S) and incubated at 55 °C for 1 h followed by 95 °C for 10 min. The equivalent volume accounting for 10 µg of protein was resolved on a 4–12% Criterion XT Bis-Tris gel (Bio-Rad 3450124) at 100 V for 1 h 45 min and LPS was detected by silver stain as described previously[40]. Briefly, the gel was fixed by incubation in 200 ml of 40% ethanol, 5% acetic acid overnight. Periodic acid was added to 0.7% and allowed to incubate for 5

min. The gel was subsequently washed with three changes of 200 ml ultrapure water over the course of 2 h. The gel was impregnated with 150 ml of freshly made staining solution (0.018 N NaOH, 0.4% NH_4OH and 0.667% silver nitrate) for 10 min and subsequently washed with three changes of 200 ml ultrapure water over the course of 2 h. The gel was developed with 200 ml freshly prepared developer solution (0.26 mM citric acid pH 3.0, 0.014% formaldehyde) and development stopped in 100 ml of 1% acetic acid. The gel was imaged in a Bio-Rad ChemiDoc MP Imaging System.

Microscopy

Overnight cultures were diluted to an OD_{600nm} of 0.01 in 3 ml LB with 1 mM IPTG as appropriate. Cultures were grown at 37 °C for 2.5 h and, when appropriate, antibiotics were added to the following concentration: 0.5 µg ml^{-1} CHIR-090, 64 µg ml^{-1} fosfomycin or 0.5 µg ml^{-1} CHIR-090 + 16 µg ml^{-1} fosfomycin. Cultures were grown at 37 °C for an additional 1.5 h and 1 ml was pelleted at 12,000*g* for 2 min. Cell pellets were resuspended to an OD_{600nm} of 10 in LB, spotted onto an LB + 1.5% agar pad, and imaged on a Nikon Eclipse 50i microscope equipped with a 100× objective and a Photometrics CoolSNAP HQ^2 camera. Images were uniformly edited using the Fiji image analysis platform version 2.3.0/1.53q[41]. Cell width was quantified using the MicrobeJ plugin (version 5.13I) of ImageJ using the following settings: area: 250–max, length: 3–100, width: 5–25, width variation: 0–0.25, width symmetry: 0–1, circularity, 0.01–1. Normality of each population was tested and determined to be lacking by both the D’Agostino and Pearson test and the Anderson–Darling test carried out by Prism (GraphPad Software, LLC). Statistical significance was determined by Kruskal–Wallis test carried out by Prism (GraphPad Software, LLC). Microscopy source images are available at https://doi.org/10.5281/zenodo.7455522 (ref. [42]).

Determination of LpxC product and substrate levels in vivo

Overnight cultures were diluted to an OD_{600nm} of 0.01 in 200 ml LB, grown at 37 °C for 3 h, and IPTG was added to a final concentration of 1 mM. After

incubation at 37 °C for one additional hour, cells were collected by centrifugation at 12,000*g* for 5 min. After resuspension in LB, cells were re-pelleted by centrifugation at 12,000*g* for 2 min. Pellets were then resuspended in 400 µl HPLC-grade H_2O. Subsequently, 250 µl chloroform and 500 µl methanol were added to the mixture. Samples were vortexed vigorously, placed on ice for 10 min, and then spun down at 4 °C at 21,000*g*. A 500 µl volume of the aqueous layer was collected and dried down at 30 °C overnight in a rotary evaporator. Dried material was resuspended in 100 µl ultrapure H_2O and the *Pa*LpxC product and substrate were detected by liquid chromatography–tandem mass spectrometry (LC–MS/MS) as described below.

In vitro LpxC activity assay

Purified F–*Pa*LpxC and H–MurA variants were diluted to 200 nM in binding buffer (25 mM Tris pH 8.0, 150 mM NaCl, 6% glycerol) to a final volume of 15 µl and mixtures were allowed to equilibrate to 30 °C for 10 min. The reaction was started by the addition of 15 µl of substrate mixture (25 mM Tris pH 8.0, 150 mM NaCl, 4% DMSO, 200 µM UDP-3-*O*-(*R*-3-hydroxydecanoyl)-GlcNAc (Carbosynth MU75071)) and allowed to proceed for 5 or 10 min at 30 °C at which point the reaction was stopped by the addition of 45 µl 2% acetic acid. Stopped reactions were frozen on dry ice, lyophilized, and the lyophilized material was resuspended in 60 µl H_2O. *Pa*LpxC product and substrate were detected by LC–MS as described below. The linear range of the reaction was determined as described above with purified F–*Pa*LpxC alone and reactions were stopped after 1, 2, 5, 10, 15 or 20 min (Extended Data Fig. 5b). H–*Pa*MurA(WT) and H–*Pa*MurA(C117S) alone exhibited background levels of LpxC activity, confirming that the observed stimulation of LpxC activity was not due to contamination within the protein preparation or an alternative enzymatic activity of *Pa*MurA (Extended Data Fig. 6e).

Activity assays with the *L. pneumophila* and *E. coli* LpxC–MurA pairs were carried out as described above except that concentration of H–MurA variants in the final reaction mixture was 200 nM. *L. pneumophila* reactions were stopped after 10 min and *E. coli* reactions were stopped after 1 min.

LC–MS

To quantify the activity of LpxC in vitro, the *Pa*LpxC product and substrate were detected by LC–MS on an Agilent Technologies 1200 series HPLC system in line with an Agilent 6520 Q-TOF mass spectrometer with electrospray ionization–mass spectrometry operating in negative mode. Samples were separated on a Waters Symmetry C18 column (5 µm; 3.9 mm × 150 mm) with a matching column guard using the following method: flow rate = 0.5 ml min^{-1}, 95% solvent A (H_2O, 10 mM ammonium formate) and 5% solvent B (acetonitrile) for 5 min followed by a linear gradient of solvent B from 5–80% over 20 min. Agilent MassHunter Workstation Qualitative Analysis software version B.06.00 was used for analysing the MS data. The *Pa*LpxC substrate (UDP-3-*O*-(*R*-3-hydroxydecanoyl)-*N*-acetylglucosamine) was quantified by monitoring the abundance of 776.21 *m*/*z* and resolved as a single peak, which was integrated to infer substrate concentration. The *Pa*LpxC product (UDP-3-*O*-(*R*-3-hydroxydecanoyl)-glucosamine) was quantified by monitoring the abundance of 734.1944 *m*/*z* and often resolved as multiple peaks as reported previously[43], all of which were integrated to infer product concentration.

*Pa*LpxC substrate and product from the aqueous fraction of methanol–chloroform-extracted whole-cell lysates were analysed by LC–MS/MS using the same settings as the LC–MS analysis described above with the following adaptations. Parent ions with *m*/*z* of 776.1986 (corresponding to the *Pa*LpxC substrate) and *m*/*z* of 734.1872 (corresponding to the *Pa*LpxC product) were targeted for MS/MS with a collision energy of 40. Fragment ions between 50 and 850 *m*/*z* were analysed. The relative abundance of the *Pa*LpxC substrate and product were quantified by integrating the peaks observed for 776.1986 *m*/*z* and 734.1872 *m*/*z*, respectively. Raw LC–MS/MS data are available at https://doi.org/10.5281/zenodo.7455522 (ref. [42]).

In vitro MurA activity assay

Purified MurA variants were diluted to 200 nM (for *Pa*MurA(WT), *Pa*MurA(G58D) and *Pa*MurA(E406K)) or 2 µM (for *Pa*MurA* variants) in MurA reaction buffer (25 mM Tris pH 7.75, 6% glycerol) to a final volume of 25 µl and allowed to equilibrate to room temperature for 30 min. The

reaction was started by the addition of 25 µl of MurA substrate mixture (25 mM Tris pH 7.75, 2 mM UDP-*N*-acetylglucosamine (Sigma-Aldrich U4375), and 1 mM phosphoenolpyruvate (Sigma-Aldrich 10108294001)) and the reaction was allowed to proceed for 30 min at 37 °C. The reaction was stopped by the addition of 800 µl Lanzetta reagent (0.033% malachite green, 1% ammonium molybdate, 1 N HCl, 0.2% Triton X-100)[44] and was incubated at room temperature for 1.25 min before the addition of 100 µl 34% sodium citrate. After an additional 30-min incubation at room temperature, the absorbance at 600 nm (A_{660nm}) of each sample was determined. A standard curve made from NaH_2PO_4 diluted in ultrapure water was used to quantify the amount of free P_i released by the reaction.

MurA–LpxC complex structure prediction

The structure of the LpxC–MurA complex was predicted using the default parameters of AlphaFold[19]. The *P. aeruginosa* PAO1 LpxC and MurA sequences or the *E. coli* MG1655 LpxC and MurA sequences were separated by a linker containing 15 repeats of Gly-Gly-Gly-Ser, which was not displayed for clarity. The structure was visualized using ChimeraX version 1.1.1 (ref. [45]).

Predicting complex interactions

We sought to use evolutionary covariation to predict the organisms in which the LpxC–MurA interaction does or does not exist. We used EVComplex run with EVcouplings v0.0.5 (ref. [18]) informed by the AlphaFold model to define an interaction score based on the inferred pairwise interaction terms \({J}_{{ij}}\) of the complex and validate our proposed complex structure.

We first generated multiple sequence alignments of LpxC and MurA using jackhammer v3.1b2 (ref. [46]) with five iterations on the Uniref100 dataset[47] released in February 2021. Both *P. aeruginosa* and *E. coli* MurA (Uniprot: Q9HVW7 and P0A749) and LpxC (Uniprot: P47205 and P0A725) were used as initial sequences. In each case, alignments were constructed with a range of sequence score thresholds from 0.1 to 0.8 bits per residue, which did not meaningfully change which sequences were included for either

species or protein. The same sequences were collected when starting from either species. For all analyses, we began with alignments of 54,940 MurA and 23,634 LpxC sequences generated from a relative bitscore of 0.8, and then concatenated MurA and LpxC sequences for each species, resolving ambiguities in pairing by selecting the sequence in each species closest to the original *P. aeruginosa* homologue. The final concatenated alignment contained 11,834 sequences. From this concatenated multiple sequence alignment, evolutionary couplings were determined using pseudo-likelihood maximization[48].

We more closely analysed the top five ranked intermolecular couplings (J_{ij}) of the direct coupling analysis model that correspond to pairs of amino acid positions in structural contact in the AlphaFold2 model, defined as having Cα atoms within 9 Å. Each pair of sequences can be scored on the coupling between MurA position *i* and LpxC position *j* on the basis of the appropriate \({J}_{{ij}}\) term from the model. Summing these five scores for each species, we observed a bimodal distribution (Extended Data Fig. 9c). We selected only pairs of sequences whose sum over the five scores exceeded 0.1, totalling 1,689 sequence pairs (Extended Data Fig. 9c), and trained a new EVcomplex model on just these sequences from the concatenated alignment. In this model, five of the top six intermolecular couplings corresponded to predicted contacts in the AlphaFold2 structure (Extended Data Fig. 9d). We defined a final 'interaction score' based on the top 20 coupling terms from this refined model and used it to score all 11,834 pairs of sequences in the concatenated alignment again. The sequence analysis files are available at https://doi.org/10.5281/zenodo.7455522 (ref. [42]). The top 20 terms were chosen arbitrarily, but this exact choice does not meaningfully affect downstream analyses, and all conclusions about relative interaction scores remain true for choices of top couplings between five and five hundred.

To estimate the percentage of interacting sequences in each clade, we fitted a two-component Gaussian mixture model using sklearn v1.0.2 (ref. [49]) to the interaction scores for all sequence pairs and calculated the posterior probability of each sequence being drawn from the upper distribution. We mapped the taxonomic identifier for each sequence pair to its class in the National Center for Biotechnology Information taxonomy database using

ETE3 v1.3.2 (ref. [50]), identifying 2,459 alphaproteobacteria and 2,781 gammaproteobacteria in the alignment. We then estimated the expected fraction of interacting sequences in each clade as the average of the interaction probabilities of each sequence in that clade.

To visualize the phylogenetic structure of these sequences, we built a maximum-likelihood phylogeny of MurA and LpxC from the concatenated multiple sequence alignment with FastTree version 2.1.11 (refs. [51]). Interaction scores were mapped onto the tree in Python v3.8.8 and visualized with the interactive tree of life (ITOL) v5 (ref. [52]). Custom code used in this study can be found at https://github.com/samberry19/evcomplex-interaction-scoring or https://doi.org/10.5281/zenodo.7471436 (ref. [53]).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

LC–MS/MS source data, microscopy images and sequence analysis files used to derive LpxC–MurA interaction scores are available at https://doi.org/10.5281/zenodo.7455522. Uniprot accession codes for genes used to generate LpxC–MurA interaction scores are included in the Methods and source data. All bacterial strains and plasmids developed in this study are available upon request. Source data are provided with this paper.

Code availability

All code generated in this study is available at https://github.com/samberry19/evcomplex-interaction-scoring or https://doi.org/10.5281/zenodo.7471436.

References

1. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414 (2010).

2. Zgurskaya, H. I., Lopez, C. A. & Gnanakaran, S. Permeability barrier of Gram-negative cell envelopes and approaches to bypass it. *ACS Infect. Dis.* **1**, 512–522 (2015).

3. Bush, K. Antimicrobial agents targeting bacterial cell walls and cell membranes. *Rev. Sci. Tech.* **31**, 43056 (2012).

4. Ogura, T. et al. Balanced biosynthesis of major membrane components through regulated degradation of the committed enzyme of lipid A biosynthesis by the AAA protease FtsH (HflB) in *Escherichia coli*. *Molec. Microbiol.* **31**, 833–844 (1999).

5. Klein, G., Kobylak, N., Lindner, B., Stupak, A. & Raina, S. Assembly of lipopolysaccharide in *Escherichia coli* requires the essential LapB heat shock protein. *J. Biol. Chem.* **289**, 14829–14853 (2014).

6. Raetz, C. R. Biochemistry of endotoxins. *Annu. Rev. Biochem.* **59**, 129–170 (1990).

7. Anderson, M. S., Bulawa, C. E. & Raetz, C. R. The biosynthesis of gram-negative endotoxin. Formation of lipid A precursors from UDP-GlcNAc in extracts of *Escherichia coli*. *J. Biol. Chem.* **260**, 15536–15541 (1985).

8. Guest, R. L., Same Guerra, D., Wissler, M., Grimm, J. & Silhavy, T. J. YejM modulates activity of the YciM/FtsH protease complex to prevent lethal accumulation of lipopolysaccharide. *mBio* **11**, e00598-20 (2020).

9. Anderson, M. S. et al. UDP-*N*-acetylglucosamine acetyltransferase of *Escherichia coli*: the first step in endotoxin biosynthesis is thermodynamically unfavorable. *J. Biol. Chem.* **268**, 19858–19865 (1993).

10. Langklotz, S., Schakermann, M. & Narberhaus, F. Control of lipopolysaccharide biosynthesis by FtsH-mediated proteolysis of LpxC

is conserved in enterobacteria but not in all gram-negative bacteria. *J. Bacteriol.* **193**, 1090–1097 (2011).

11. Kahan, F. M., Kahan, J. S., Cassidy, P. J. & Kropp, H. The mechanism of action of fosfomycin (phosphonomycin). *Ann. N. Y. Acad. Sci.* **235**, 364–386 (1974).

12. Strominger, J. L. Enzymic transfer of pyruvate to uridine diphosphoacetylglucosamine. *Biochem. Biophys. Acta* **30**, 645–646 (1958).

13. Zhu, J. Y. et al. Functional consequence of covalent reaction of phosphoenolpyruvate with UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA). *J. Biol. Chem.* **287**, 12657–12667 (2012).

14. Eschenberg, S. & Schonbrunn, E. Comparative X-ray analysis of the un-liganded fosfomycin-target MurA. *Proteins* **40**, 290–298 (2000).

15. Wanke, C. & Amrhein, N. Evidence that the reaction of the uDP-N-acetylglucosamine 1-carboxyvinyltransferase proceeds through the O-phosphothioketal of pyruvic acid bound to Cys115 of the enzyme. *Eur. J. Biochem.* **218**, 861–870 (1993).

16. Eschenburg, S., Priestman, M. & Schönbrunn, E. Evidence that the fosfomycin target Cys115 in UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) is essential for product release. *J. Biol. Chem.* **280**, 3757–3763 (2005).

17. Schönbrunn, E., Eschenberg, S., Krekel, F., Luger, K. & Amrhein, N. Role of the loop containing residue 115 in the induced-fit mechanism of the bacterial cell wall enzyme MurA. *Biochemistry* **39**, 2164–2173 (2000).

18. Hopf, T. A. et al. Sequence co-evolution gives 3D contacts and structures of protein complexes. *Elife* **3**, e03430 (2014).

19. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

20. Mohan, S., Kelly, T. M., Eveland, S. S., Raetz, C. R. & Anderson, M. S. An *Escherichia coli* gene (FabZ) encoding (3R)-hydroxymyristoyl acyl carrier protein dehydrase. Relation to fabA and suppression of mutations in lipid A biosynthesis. *J. Biol. Chem.* **269**, 32896–32903 (1994).

21. Thomanek, N. et al. Intricate crosstalk between lipopolysaccharide, phospholipid and fatty acid metabolism in *Escherichia coli* modulates proteolysis of LpxC. *Front. Microbiol.* **9**, 3285 (2018).

22. Mostafavi, M. et al. Interplay of *Klebsiella pneumoniae* fabZ and lpxC mutations leads to LpxC inhibitor-dependent growth resulting from loss of membrane homeostasis. *mSphere* **3**, e00508-18 (2018).

23. Zeng, D. et al. Mutants resistant to LpxC inhibitors by rebalancing cellular homeostasis. *J. Biol. Chem.* **288**, 5475–5486 (2013).

24. Mizyed, S. et al. UDP-*N*-acetylmuramic acid (UDP-MurNAc) is a potent inhibitor of MurA (enolpyruvyl-UDP-GlcNAc synthase). *Biochemistry* **44**, 4011–4017 (2005).

25. Fivenson, E. M. & Bernhardt, T. G. An essential membrane protein modulates the proteolysis of LpxC to control lipopolysaccharide synthesis in *Escherichia coli*. *mBio* **11**, e00939-20 (2020).

26. Clairfeuille, T. et al. Structure of the essential inner membrane lipopolysaccharide–PbgA complex. *Nature* **584**, 479–483 (2020).

27. Collier, D. N., Hager, P. W. & Phibbs, P. V. Jr Catabolite repression control in the *Pseudomonads*. *Res. Microbiol.* **147**, 551–561 (1996).

28. Kline, T. Potent, novel in vitro inhibitors of the *Pseudomonas aeruginosa* deacetylase LpxC. *J. Med. Chem.* **45**, 3112–3129 (2002).

29. McClerren, A. L. et al. A slow, tight-binding inhibitor of the zinc-dependent deacetylase LpxC of Lipid A biosynthesis with antibiotic activity comparable to ciprofloxacin. *Biochemistry* **44**, 16574–16583 (2005).

30. Langsdorf, E. F. et al. Screening for antibacterial inhibitors of the UDP-3-*O*-(*R*-3-hydroxymyristoyl)-*N*-acetylglucosamine deacetylase (LpxC) using a high-throughput mass spectrometry assay. *J. Biomol. Screen.* **15**, 52–61 (2010).

31. Erwin, A. L. Antibacterial drug discovery targeting the lipopolysaccharide biosynthetic enzyme LpxC. *Cold Spring Harb. Perspect. Med.* **6**, a025304 (2016).

32. Krause, K. M. et al. Potent LpxC inhibitors with *in vitro* activity against multidrug-resistant *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **63**, e00977-19 (2019).

33. Ushiyama, F. et al. Lead optimization of 2-hydroxymethyl imidazoles as non-hydroxamate LpxC inhibitors: discovery of TP0586532. *Bioorg. Med. Chem.* **30**, 115964 (2021).

34. Hernick, M. & Fierke, C. A. Catalytic mechanism and molecular recognition of *E. coli* UDP-3-*O*-(*R*-3-hydroxymyristoyl)-*N*-acetylglucosamine deacetylase probed by mutagenesis. *Biochemistry* **45**, 15240–15248 (2006).

35. Miller, J. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, 1972).

36. Yang, D. C., Tan, K., Joachimiak, A. & Bernhardt, T. G. A conformational switch controls cell wall-remodelling enzymes required for bacterial cell division. *Mol. Microbiol.* **85**, 768–781 (2012).

37. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858 (1996).

38. Peng, J. & Gygi, S. P. Proteomics: the move to mixtures. *J. Mass Spectrom.* **36**, 1083–1091 (2001).

39. Eng, J. K., McCormac, A. L. & Yates, J. R. III An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).

40. Tsai, C.-M. & Frasch, C. E. A sensitive silver stain for detecting lipopolysaccharide in polyacrylamide gels. *Anal. Biochem.* **119**, 115–119 (1982).

41. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

42. Hummels, K. R. et al. Data for "Coordination of bacterial cell wall and outer membrane biosynthesis". Zenodo https://doi.org/10.5281/zenodo.7455522 (2022).

43. Cohen, F. et al. Optimization of LpxC inhibitors for antibacterial activity and cardiovascular safety. *ChemMedChem* **14**, 1560–1572 (2019).

44. Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. An improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.* **100**, 95–97 (1979).

45. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

46. Johnson, L. S., Eddy, S. R. & Portugaly, E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinform.* **11**, 431 (2010).

47. Suzek, B. E. et al. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926–932 (2015).

48. Ekeberg, M., Lövkvist, C., Lan, Y., Weigt, M. & Aurell, E. Improved contact prediction in proteins: using pseudolikelihoods to inver Potts models. *Phys. Rev. E* **87**, 012707 (2013).

49. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

50. Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–

1638 (2016).

51. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009).

52. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

53. Berry, S. P. samberry19/evcomplex-interaction-scoring: v1.0 (v1.0). Zenodo https://doi.org/10.5281/zenodo.7471436 (2022).

Acknowledgements

We thank all of the members of the Bernhardt, Rudner and Mekalanos labs for their thoughtful discussions and advice throughout this project, especially L. Marmont for technical insight regarding protein purification. We also thank S. Lory for strains and reagents as well as advice on the genetic manipulation of *P. aeruginosa*, B. Silva and M. Welsh for assistance with LC–MS/MS analysis, and the Taplin Mass Spectrometry Core Facility at Harvard Medical School for the identification of co-affinity-purified proteins. LC–MS data were acquired on an Agilent 6520 Q-TOF spectrometer supported by the Taplin Funds for Discovery Program. Protein structures were visualized with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. This work was supported by the National Institutes of Health (F32AI164630 to K.R.H., AI148752 to S.W., AI083365 to T.G.B., and U19 AI158028 to T.G.B and S.W.), the Chan Zuckerberg Initiative (CZI2018-191853 to D.S.M.) and Investigator funds from the Howard Hughes Medical Institute (T.G.B).

Author information

Authors and Affiliations

1. Department of Microbiology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA

   Katherine R. Hummels, Zhaoqi Li, Atsushi Taguchi, Suzanne Walker & Thomas G. Bernhardt

2. Department of Systems Biology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA

   Samuel P. Berry, Joseph K. Min & Debora S. Marks

3. SANKEN (The Institute of Scientific and Industrial Research), Osaka University, Ibaraki, Japan

   Atsushi Taguchi

4. Howard Hughes Medical Institute, Boston, MA, USA

   Thomas G. Bernhardt

Contributions

K.R.H. and T.G.B. designed the study and wrote the manuscript. S.P.B., J.K.M. and D.S.M. devised and carried out the covariation analysis. Z.L. established the methanol–chloroform metabolite extraction protocol. A.T. and S.W. developed the LC–MS protocol. K.R.H. carried out all other experiments and data analysis. All authors read and approved the manuscript.

Corresponding author

Correspondence to Thomas G. Bernhardt.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Russell Bishop and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Regulation of LpxC in *P. aeruginosa* differs from that observed in *E. coli*.

(a) Anti-His immunoblot detecting H-^{Pa}LpxC expressed from the native chromosomal locus in wild-type cells or an *ftsH* deletion mutant. A corresponding blot for RpoA was used as a loading control. Data are representative of 3 biological replicates. (b) Spot titer assay in which serial dilutions of PAO1 harboring an empty plasmid or one encoding *His-*^{Pa}*lpxC* or *His-*^{Ec}*lpxC* under arabinose-inducible control were plated on LB agar supplemented with arabinose and/or the LpxC inhibitor CHIR-090 as indicated. Plates were incubated at 37 °C for 20 h before being imaged. Data are representative of 3 biological replicates. (c) Anti-His immunoblot analysis of His-^{Pa}LpxC or His-^{Ec}LpxC protein levels in exponentially growing PAO1. Immunoblot for RpoA serves as a loading control. Data are representative of 3 biological replicates. For gel source data, see Supplementary Fig. 1.

Extended Data Fig. 2 LPS levels are altered upon mis-regulation of LpxC.

Silver stain of LPS harvested from exponentially growing cultures and western blot of RpoA from the same samples as a loading control. (a) PAO1 harboring an empty plasmid or one encoding ^{Pa}*lpxC* or ^{Ec}*lpxC* under arabinose-inducible control were induced with 0.1% arabinose 1 h prior to harvesting samples. (b) PAO1, PA1118 [Δ*murA* P_{lac}-*murA*], or PA1135 [Δ*murB* P_{lac}-*murB*] were grown in the presence or absence of 1 mM IPTG as indicated before samples were processed. (C) PAO1 harboring an empty plasmid or one encoding ^{Pa}*murA(WT)* or ^{Pa}*murA(C117S)* under IPTG-inducible control were induced with 1 mM IPTG 1 h prior to harvesting samples. Data are representative of 3 biological replicates. For gel source data, see Supplementary Fig. 1.

Extended Data Fig. 3 MurA is essential and its depletion phenocopies simultaneous inhibition of PG and LPS biosynthesis.

(a–c) Growth curves of *P. aeruginosa* strains in LB with or without IPTG as indicated. Dots represent the average of 3 biological replicates and dashed lines indicate the standard deviation. The following strains were used: (a) PAO1 [WT] and PA1080 [Δ*PA4701*], (b,c) PAO1 [WT], PA1118 [Δ*murA* P_{lac}-*murA*], and PA1135 [Δ*murB* P_{lac}-*murB*]. (d) Phase contrast images of *P. aeruginosa* cells after 1 hr treatment with the indicated antibiotic(s). Scale bar indicates 2 µm. Data are representative of at least 2 biological replicates (e) Phase contrast images of the indicated *P. aeruginosa murA* or *murB* depletion strains grown for 4 h in the presence or absence of inducer as indicated. MurB depletion was analyzed as a control to compare the phenotype of inactivating another early step in PG synthesis with that of MurA. Scale bar indicates 2 µm. Data are representative of at least 3 biological replicates (f) Quantification of cell width after 1 hr treatment with the indicated antibiotic(s) or after depletion of MurA or MurB. Each dot represents an individual cell and the median of the population is indicated by a black line. n indicates the number of cells analyzed. For each condition,

the cells quantified were derived from a single population and data are representative of biological duplicates.

Source Data

Extended Data Fig. 4 ^{Pa}MurA interacts with ^{Pa}LpxC.

(a) H-^{Pa}LpxC *in vivo* pulldowns. The expression status of H-^{Pa}LpxC and F-^{Pa}MurA before (input) and after (elution) co-affinity purification using Ni-NTA resin is indicated above the immunoblots. The variant of F-^{Pa}MurA produced is indicated by WT for F-^{Pa}MurA(WT) or * for F-^{Pa}MurA(C117S). When indicated, 0.5 µg/mL CHIR-090 was added to cultures 1 hr prior to harvesting and was maintained in all lysis and wash buffers as detailed in the methods section. The following strains were used to generate the lysates: PA239 (lane 1), PA1013 (lane 2), PA1068 (lane 3), PA1071 (lanes 4 and 6), and PA1121 (lanes 5 and 7). Data are representative of at 3 biological replicates. (b) Purified F-^{Pa}LpxC (2.5 µM) was mixed in a 1:1 ratio with purified H-MurA variants in the presence or absence of CHIR-090 (5.7 µM) as indicated. The mixtures were pulled down with anti-FLAG resin and the input and elution subjected to SDS-PAGE and Coomassie staining. Mobilities of H-MurA and F-LpxC are indicated by a cyan or gold carrot, respectively. Data are representative of 3 replicates. For gel source data, see Supplementary Fig. 1.

Extended Data Fig. 5 Purified proteins used in this study and linearity of LpxC activity assay.

(a) Purified proteins used in this study were resolved by SDS-PAGE and protein was visualized with Coomassie staining. Data are representative of at least two 2 replicates. For gel source data, see Supplementary Fig. 1. (b) Time course in which turnover of UDP-3-O-(R-3-hydroxydecanoyl)-N-acetylglucosamine (^{Pa}LpxC substrate) to UDP-3-O-(R-3-hydroxydecanoyl)-glucosamine (^{Pa}LpxC product) by FLAG-^{Pa}LpxC over the course of 20 min was monitored by LC-MS. The R^2 value of the linear regression is presented.

Source Data

Extended Data Fig. 6 Mutations in the ^{Pa}MurA active site are toxic but can be suppressed by inhibition of ^{Pa}LpxC.

(a) Spot titer assay in which serial dilutions of PAO1 harboring an empty vector, one encoding ^{Pa}MurA(WT) or the indicated ^{Pa}MurA variant were plated on LB agar supplemented with IPTG and/or CHIR-090 as indicated. Plates were incubated at 37 °C for 20 h before imaging. † indicates the presence of a silent mutation in the construct. See Table S2 for details. Data are representative of 3 biological replicates. (b) Crystal structure of *E. cloacae* MurA (**PDB** 1EJC**)**[14] in which residues corresponding to the identified ^{Pa}MurA dominant negative alleles are depicted in gold spheres, or in the case of Cys117, a red sphere. Note that the substitutions all cluster around the active site. (c) MurA activity assay in which purified ^{Pa}MurA variants (100 nM) were mixed with UDP-GlcNAc (1 mM) and PEP (0.5 mM) and the release of P_i was measured by Lanzetta assay. Dots indicate the values obtained for three individual replicates, bars indicate the mean, and error bars represent their standard deviation. (d) Catalytic activity of purified ^{Pa}LpxC (100 nM) alone or in the presence of MurA variants (100 nM) assayed by conversion of UDP-3-O-(R-3-hydroxydecanoyl)-N-acetylglucosamine (^{Pa}LpxC substrate) to UDP-3-O-(R-3-hydroxydecanoyl)-glucosamine (^{Pa}LpxC product). Dots indicate the values obtained for three individual replicates, bars indicate the mean, and error bars represent their standard deviation. (e) LpxC enzymatic activity detected from preparation of MurA variants (100 nM) alone assayed as in panel d. Dots indicate the values obtained for three individual replicates, bars indicate the mean, and error bars represent their standard deviation.

Source Data

Extended Data Fig. 7 LC-MS/MS analysis of ^{Pa}LpxC substrate and product.

(a) Chemical structure of the ^{Pa}LpxC substrate. The acetyl moiety removed by LpxC is highlighted in gold. Boxed regions indicate putative fragment ions highlighted in panels c, d, and h-j. (b) Extracted ion chromatogram (EIC) of UDP-3-O-(R-3-hydroxydecanoyl)-N-acetylglucosamine (^{Pa}LpxC substrate, 776.1986 *m/z*, black line) and UDP-3-O-(R-3-hydroxydecanoyl)-glucosamine (^{Pa}LpxC product, 734.1872 *m/z*, gold line) derived from *in vitro* reaction mixtures containing the ^{Pa}LpxC substrate along with purified ^{Pa}LpxC and ^{Pa}MurA resolved using LC-MS operating in negative mode. The dashed lines indicate the peaks assigned to the ^{Pa}LpxC substrate (S) and ^{Pa}LpxC product (P). (c—d) MS/MS spectrum associated with the panel b peaks S and P, respectively. The parent ion is indicated by an asterisk and fragment ions corresponding to those highlighted in panel a are labeled. (e-g) EICs ^{Pa}LpxC product and ^{Pa}LpxC substrate detected by LC-MS in the aqueous fraction of methanol-chloroform extracted whole cell lysates. The dashed lines indicate the peak assigned to the ^{Pa}LpxC substrate (S) and ^{Pa}LpxC product peaks (P_1 and P_2). Note that the ^{Pa}LpxC product has been previously reported to resolve as two peaks[43], both of which were integrated to infer the relative product abundance. Data are representative of biological triplicates. (h-j) MS/MS spectrum associated with the panel g peaks S, P_1, and P_2, respectively. The parent ion is indicated by an asterisk and fragment ions corresponding to those highlighted in panel a are labeled.

Source Data

Extended Data Fig. 8 MurA(G58D) and MurA(E406K) impact binding and activation of ^{Pa}LpxC.

(a) Anti-FLAG immunoblot detecting F-^{Pa}MurA variants after 1 h of induction with 1 mM IPTG. A corresponding blot for RpoA was used as a loading control. Data are representative of 3 biological replicates. (b) MurA activity assay in which purified ^{Pa}MurA variants (100 nM) were mixed with UDP-GlcNAc (1 mM) and PEP (0.5 mM) and the release of P_i was measured by Lanzetta assay. Dots indicate the values obtained for three individual replicates, bars indicate the mean, and error bars represent their standard deviation. The dashed line indicates the average catalytic activity of

^{Pa}MurA(C117S) observed in Fig S7C. (c) *in vitro* pulldowns in which purified F-^{Pa}LpxC and H-MurA variants were mixed in a 1:1 ratio in the presence or absence of CHIR-090 and processed as in Extended Data Fig. 4b. Data are representative of at least two replicates. (d) Spot titer assay in which serial dilutions of a PAO1 strain harboring a ^{Pa}*murA* deletion or ^{Pa}*murA(C117S)* allele at the native locus complemented by a chromosomally-integrated, IPTG-inducible copy of ^{Pa}*murA(WT)* were plated on LB agar with the indicated supplements. As indicated, the strains also contained an empty plasmid or one encoding ^{Pa}*murA(G58D)* under arabinose-inducible control. Plates were incubated at 37°C for 20 h before being photographed. Data are representative of 3 biological replicates. For gel source data, see Supplementary Fig. 1.

Source Data

Extended Data Fig. 9 ^{Pa}LpxC and ^{Pa}MurA are predicted to interact.

(a,b) AlphaFold2 predicted aligned error matrices of the ^{Pa}LpxC/^{Pa}MurA complex and ^{Ec}LpxC/^{Ec}MurA complex structures, respectively. (c) Distribution of initial interaction scores among all LpxC and MurA pairs analyzed. Red bars indicate the sequences used to train the final EVcomplex model and the score corresponding to the ^{Pa}LpxC/^{Pa}MurA pair is indicated by a dashed line. (d) Model structure of the ^{Pa}LpxC/^{Pa}MurA complex predicted by AlphaFold2[19]. ^{Pa}LpxC is represented in gold and ^{Pa}MurA is represented in cyan. The top six intermolecular couplings between LpxC/MurA residues from the final EVcomplex model are highlighted in red with red lines connecting the coupling pairs. (e) Distribution of final interaction scores among all LpxC and MurA pairs analyzed. The data fit a two-component Gaussian mixture model (black line) indicating the presence of a population with low interaction scores (blue line) and high interaction scores (orange line). (f) Catalytic activity of purified ^{Lpn}LpxC (100 nM) alone or in the presence of ^{Lpn}MurA (200 nM) assayed by conversion of UDP-3-O-(R-3-hydroxydecanoyl)-N-acetylglucosamine to UDP-3-O-(R-3-hydroxydecanoyl)-glucosamine. Dots indicate the values obtained for three individual replicates, bars indicate the mean, and error bars represent their

standard deviation. (g) Catalytic activity of purified ^{Ec}LpxC (100 nM) alone or in the presence of ^{Ec}MurA (200 nM) assayed by conversion of UDP-3-O-(R-3-hydroxydecanoyl)-N-acetylglucosamine to UDP-3-O-(R-3-hydroxydecanoyl)-glucosamine. Dots indicate the values obtained for three individual replicates, bars indicate the mean, and error bars represent their standard deviation.

Source Data

Supplementary information

Supplementary Information

This document contains Supplementary Methods, Tables 1–5, and Supplementary References.

Reporting Summary

Peer Review File

Supplementary Fig. 1

Original source images for all electrophoresis gels reported in this study.

Supplementary Data

Source data for Supplementary Table 1.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Source Data Extended Data Fig. 9

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Hummels, K.R., Berry, S.P., Li, Z. *et al.* Coordination of bacterial cell wall and outer membrane biosynthesis. *Nature* **615**, 300–304 (2023).

https://doi.org/10.1038/s41586-023-05750-0

- Received: 16 July 2021
- Accepted: 23 January 2023
- Published: 01 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05750-0

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

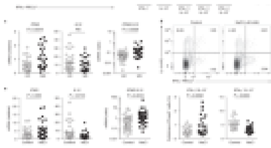
Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

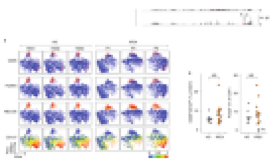
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05750-0



Activated β-catenin in Foxp3+ regulatory T cells links inflammatory environments to autoimmunity

29 October 2018

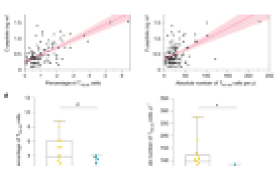
Tomokazu Sumida, Matthew R. Lincoln, … David A. Hafler



Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic CD4+ T cell perturbations

08 April 2021

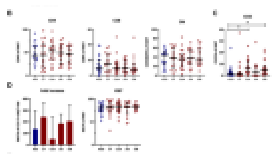
David Zemmour, Louis-Marie Charbonnier, … Christophe Benoist



Type 1 diabetes progression is associated with loss of CD3+CD56+ regulatory T cells that control CD8+ T-cell effector functions

17 February 2020

Giuseppe Terrazzano, Sara Bruzzaniti, … Mario Galgani



[**Immunometabolic factors in adolescent chronic disease are associated with Th1 skewing of invariant Natural Killer T cells**]

11 October 2021

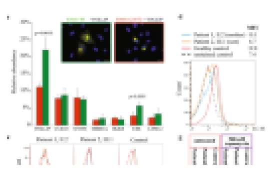
Francesca A. Ververs, Suzanne E. Engelen, … Henk S. Schipper



[**Elevated T cell levels in peripheral blood predict poor clinical response following rituximab treatment in new-onset type 1 diabetes**]

21 June 2018

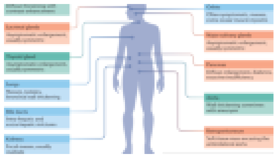
Peter S. Linsley, Carla J. Greenbaum, … Matthew J. Dufort



[**Dominant TOM1 mutation associated with combined immunodeficiency and autoimmune disease**]

27 June 2019

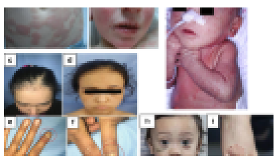
Salla Keskitalo, Emma M. Haapaniemi, … Markku Varjosalo



[IgG4-related disease: an update on pathophysiology and implications for clinical care](#)

16 September 2020

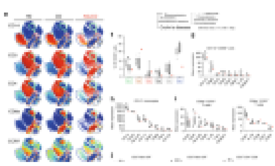
Cory A. Perugino & John H. Stone



[Treg-associated monogenic autoimmune disorders and gut microbial dysbiosis](#)

17 March 2021

Yuying Liu, Jasmin Freeborn, … Jon Marc Rhoads



[Leptin induces TNFα-dependent inflammation in acquired generalized lipodystrophy and combined Crohn's disease](#)

10 December 2019

Jörn F. Ziegler, Chotima Böttcher, … Carl Weidinger

- Article
- Published: 22 February 2023

Autoimmunity in Down's syndrome via cytokines, CD4 T cells and CD11c^{+} B cells

- Louise Malle,
- Roosheel S. Patel,
- Marta Martin-Fernandez,
- O Jay Stewart,
- Quentin Philippot,
- Sofija Buta,
- Ashley Richardson,
- Vanessa Barcessat,
- Justin Taft,
- Paul Bastard,
- Julie Samuels,
- Clotilde Mircher,
- Anne-Sophie Rebillat,
- Louise Maillebouis,
- Marie Vilaire-Meunier,
- Kevin Tuballes,
- Brad R. Rosenberg,
- Rebecca Trachtman,
- Jean-Laurent Casanova,
- Luigi D. Notarangelo,
- Sacha Gnjatic,
- Douglas Bush &
- …
- Dusan Bogunovic

Show authors

Nature **volume 615**, pages 305–314 (2023)

- 13k Accesses

- 238 Altmetric
- Metrics details

Abstract

Down's syndrome (DS) presents with a constellation of cardiac, neurocognitive and growth impairments. Individuals with DS are also prone to severe infections and autoimmunity including thyroiditis, type 1 diabetes, coeliac disease and alopecia areata[1,2]. Here, to investigate the mechanisms underlying autoimmune susceptibility, we mapped the soluble and cellular immune landscape of individuals with DS. We found a persistent elevation of up to 22 cytokines at steady state (at levels often exceeding those in patients with acute infection) and detected basal cellular activation: chronic IL-6 signalling in CD4 T cells and a high proportion of plasmablasts and $CD11c^{+}Tbet^{high}CD21^{low}$ B cells (Tbet is also known as TBX21). This subset is known to be autoimmune-prone and displayed even greater autoreactive features in DS including receptors with fewer non-reference nucleotides and higher *IGHV4-34* utilization. In vitro, incubation of naive B cells in the plasma of individuals with DS or with IL-6-activated T cells resulted in increased plasmablast differentiation compared with control plasma or unstimulated T cells, respectively. Finally, we detected 365 auto-antibodies in the plasma of individuals with DS, which targeted the gastrointestinal tract, the pancreas, the thyroid, the central nervous system, and the immune system itself. Together, these data point to an autoimmunity-prone state in DS, in which a steady-state cytokinopathy, hyperactivated CD4 T cells and ongoing B cell activation all contribute to a breach in immune tolerance. Our findings also open therapeutic paths, as we demonstrate that T cell activation is resolved not only with broad immunosuppressants such as Jak inhibitors, but also with the more tailored approach of IL-6 inhibition.

Main

First desc ribed by John Langdon Down in 1866, Down's syndrome (DS) or trisomy 21 is the most common chromosomal anomaly in the USA today, affecting 1 in 700 newborn babies[3,4]. This extra copy of around 200 genes

results in a syndrome with considerable phenotypic variability that includes intellectual disability, developmental malformations—particularly of the heart and the gut—and increased risk of Alzheimer's disease[1]. As care for individuals with DS has substantially improved in recent decades[5], the immune features of DS have become apparent: patients have an increased risk of severe infectious disease concomitant with a higher incidence of autoimmunity including thyroiditis (50%), coeliac disease (5%), alopecia areata (1–11%) and type 1 diabetes (1%)[1,2,6].

Recent studies into the molecular mechanisms of immunological disease have focused on the overactive interferon (IFN) response[7,8,9] and thymic dysfunction[10] reported in individuals with DS as most IFN receptor subunits and *AIRE* are expressed from chromosome 21. On the innate side, monocytes from individuals with DS exhibit basal IFN-I and II signalling and hyper-respond to IFNα and IFNγ stimulation[7]. On the adaptive side, thymic architecture perturbations[10] and T cell polarization towards differentiated subsets have been described[11]. Furthermore, low B cell counts in individuals with DS have been documented for decades[12] and recent investigations have identified decreased proliferation and increased apoptosis in this population[13].

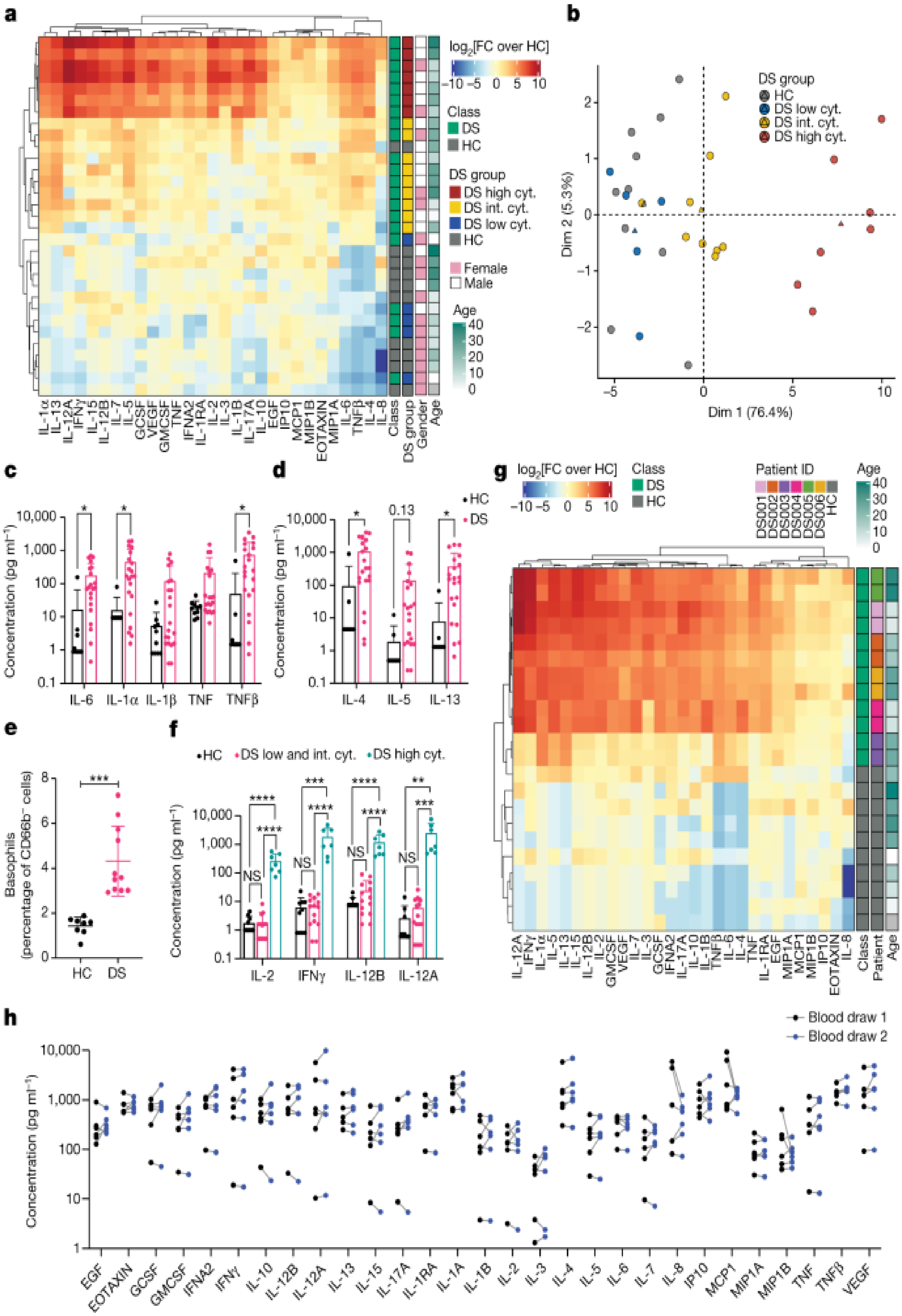
It is now understood that a tightly regulated immune response is critical to prevent infection and excessive inflammation: the absence of a well-orchestrated immune response leads to opportunistic infections, whereas an overactive immune response leads to systemic organ damage[14]. Where DS lies on this spectrum of immune dysregulation and how it contributes to clinical manifestations is still largely unknown.

Down's syndrome is a cytokinopathy

To capture the soluble immune landscape at steady-state in DS, we performed a cytokine array on plasma from individuals with DS ($n = 21$) and age-matched healthy control (HC, $n = 10$) individuals. Donors had no signs of infection at sampling. On the basis of unsupervised hierarchical clustering on 29 analytes, individuals with DS segregated into 3 distinct categories (Fig. 1a,b). One-third had widespread soluble immune dysregulation with up

to 2,000-fold elevation in 22 out of the 29 markers assayed. In 9 out of 21 individuals, a subset of cytokines was significantly elevated compared with in the control individuals (including IL-13, IL-4, TNFβ (also known as LTα) IL-6 and IL-1α). Finally, the remaining five individuals with DS clustered with HC individuals. As previously reported for HC individuals[15], there was a correlation between inflammatory cytokine profiles and age in DS (analysis of variance (ANOVA), $P = 0.0231$; Extended Data Fig. 1a). After scoring samples based on clinical immunological manifestations (Methods and Extended Data Table 1), we observed a significant association between immune scores and the cytokine-based clusters (ANOVA, $P = 0.0141$; Extended Data Fig. 1b). It is not yet clear whether dysregulated cytokines drive clinical immune dysfunction in individuals with DS, or vice versa.

Fig. 1: Cytokine profiling indicates broad immune dysregulation in most individuals with DS.


a
log2[FC over HC]
−10 −5 0 5 10
Class
DS
HC
DS group
DS high cyt.
DS int. cyt.
DS low cyt.
HC
Female
Male
Age
40
30
20
10
0
IL-1α IL-13 IL-12A IFNγ IL-15 IL-12B IL-7 IL-5 GCSF VEGF GMCSF TNF IFNA2 IL-1RA IL-2 IL-3 IL-1B IL-17A IL-10 EGF IP10 MCP1 MIP1B EOTAXIN MIP1A IL-6 TNFβ IL-4 IL-8 Class DS group Gender Age
b
DS group
HC
DS low cyt.
DS int. cyt.
DS high cyt.
Dim 2 (5.3%)
Dim 1 (76.4%)
c
Concentration (pg ml−1)
IL-6 IL-1α IL-1β TNF TNFβ
d
HC
DS
0.13
IL-4 IL-5 IL-13
e
Basophils (percentage of CD66b− cells)
HC DS
f
HC
DS low and int. cyt.
DS high cyt.
NS
IL-2 IFNγ IL-12B IL-12A
g
log2[FC over HC]
Class
DS
HC
Patient ID
DS001 DS002 DS003 DS004 DS005 DS006 HC
Age
IL-12A IFNγ IL-1α IL-5 IL-13 IL-15 IL-12B IL-2 GMCSF VEGF IL-7 IL-3 GCSF IFNA2 IL-17A IL-10 IL-1B TNFβ IL-6 IL-4 TNF IL-1RA EGF MIP1A MCP1 MIP1B IP10 EOTAXIN IL-8 Class Patient Age
h
Blood draw 1
Blood draw 2
Concentration (pg ml−1)
EGF EOTAXIN GCSF GMCSF IFNA2 IFNγ IL-10 IL-12B IL-12A IL-13 IL-15 IL-17A IL-1RA IL-1A IL-1B IL-2 IL-3 IL-4 IL-5 IL-6 IL-7 IL-8 IP10 MCP1 MIP1A MIP1B TNF TNFβ VEGF

a, Multiplex cytokine analysis using the magnetic Luminex assay of plasma from individuals with DS ($n = 21$) and HC individuals ($n = 10$), expressed as the $\log_2$-transformed fold change (FC) over the mean HC per cytokine (cyt.). Unsupervised clustering of samples and cytokines using the complete method (distance metric, Euclidean). Int., intermediate. **b**, PCA analysis of serum cytokines from the samples. **c**,**d**, Raw values of acute phase proteins (**c**) and T_H2 cytokines (**d**) in the plasma of individuals with DS ($n = 21$) and HC individuals ($n = 10$) measured using the magnetic Luminex assay. **e**, The frequency of basophils expressed as the percentage of $CD66b^-$ cells (non-granulocytes) from adults with DS ($n = 11$) and age-matched HC individuals ($n = 8$). **f**, Raw values of IL-2 and T_H1 cytokines in the plasma of individuals with DS ($n = 21$) and HC individuals ($n = 10$) measured using the magnetic Luminex assay. **g**,**h**, Multiplex cytokine analysis using the magnetic Luminex assay of plasma from blood drawn at separate timepoints from individuals with DS ($n = 6$) expressed as $\log_2$-transformed fold change over the mean HC individuals per cytokine followed by unsupervised clustering (**g**) and as raw values (**h**). No significant differences were detected on the basis of paired *t*-tests between blood draws for each individual for each cytokine ($P > 0.05$ for all pairs). For **c**–**f**, data are mean ± s.d. Significance was assessed using two-tailed unpaired *t*-tests (**c**–**e**) and ANOVA with Tukey's post hoc analysis (**f**); *$P \leq 0.05$; **$P \leq 0.005$; ***$P \leq 0.0005$; ****$P \leq 0.0001$.

To determine the magnitude of this global cytokine dysregulation, we compared the cytokine profiles of individuals with DS and HC individuals at steady state, in mild or severe COVID-19 ($n = 13$ (control) and $n = 7$ (individuals with DS)) or another acute respiratory infection ($n = 1$ individual with DS) (Extended Data Fig. 1c). The COVID-19 samples were collected during hospitalization or at follow-up. Samples from donors with DS and with COVID-19 were available at only one point, whereas control samples were collected during the acute phase and at follow-up. Notably, the cytokine profiles of the patients with an infection (irrespective of ploidy, disease severity or infecting virus) were not significantly different from those of uninfected individuals with DS—unbiased hierarchical clustering placed patients with an infection across all three groups of individuals with DS. The uninfected high-cytokine DS group had a broader, more severe inflammatory profile compared with any patient with an infection. Our

findings suggest that at least a third of individuals with DS have cytokine levels similar to severe acute infection at the baseline.

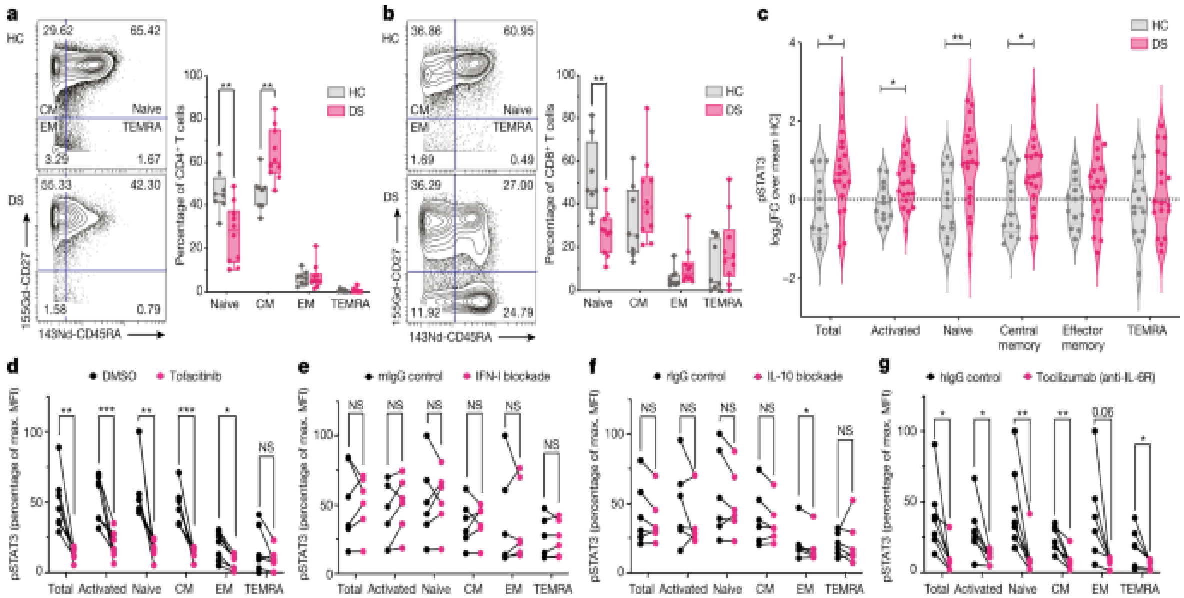
The acute-phase proteins IL-6, IL-1α and TNFβ were basally elevated in most individuals with DS (Fig. 1a,c), consistent with previous studies[16,17]. Intracellular staining suggested that $CD16^+$ monocytes, conventional dendritic cells (cDCs), and central memory CD4 and CD8 T cells from individuals with DS contained slightly more IL-6 (Extended Data Fig. 1d). T helper 2 (T_H2) cytokines IL-4 and IL-13, two central drivers of the allergic response[18], were significantly elevated in the plasma of individuals with DS (Fig. 1a,d), possibly explained by the concurrent increase in basophils (Fig. 1e). The role of other T cell subsets, especially T_H2 and T_H9 cells, must be further examined. IL-2, T_H1 cytokines IL-12 and IFNγ, and the T_H17 cytokine IL-17 were elevated in only the high-cytokine subgroup of individuals with DS (Fig. 1a,f).

Measuring cytokine levels in multiple blood samples, drawn 5 to 10 months apart, we found that the immune profile of individuals with DS was highly stable for both the high- and low-cytokine groups (Fig. 1g,h). Our findings suggest that individuals with DS have stable, long-lasting perturbations in their cytokine levels similar to those in acute COVID-19. We conclude that DS can be considered to be a cytokinopathy.

Basal IL-6-mediated T cell activation

Cytometry by time-of-flight (CyTOF)-based immunophenotyping of whole blood from individuals with DS and age-matched HC individuals (Extended Data Fig. 2a–i) revealed that both CD4 and CD8 T cells in DS were skewed towards a memory phenotype (Fig. 2a,b and Extended Data Fig. 2d,e). There were fewer naive CD4 and CD8 T cells in individuals with DS, with a concurrent increase in CD4 central memory frequency[11,19]. We also uncovered baseline phosphorylated STAT3 (pSTAT3) in naive, activated and central memory CD4 T cells of individuals with DS, suggesting active cytokine signalling (Fig. 2c and Extended Data Fig. 3a).

Fig. 2: T cell activation in DS is rescued by Jak inhibition or IL-6 blockade.



a,**b**, Representative plots and calculated frequencies of $CD4^+$ (**a**) and $CD8^+$ T cell naive, central memory (CM), effector memory (EM) and terminally differentiated effector memory (TEMRA) (**b**) subsets in whole blood from adults with DS ($n = 10$) and age-matched HC individuals ($n = 8$). **c**, Basal STAT3 phosphorylation in $CD4^+$ T cell subsets from individuals with DS ($n = 19$) and age-matched HC individuals ($n = 13$) and expressed as the $\log_2$-transformed fold change over the mean of HC individuals per subset. **d**–**g**, STAT3 phosphorylation in $CD4^+$ T cell subsets from individuals with DS, normalized to the maximum value per experiment after ex vivo whole-blood treatment for 4 h with JAK inhibition (tofacitinib) (500 nM) ($n = 7$) (**d**); IFN blockade (anti-IFNAR2 (5 μg ml^{-1}), anti-IFNα (0.2 μg ml^{-1}) and anti-IFNβ (0.2 μg ml^{-1}) antibodies) ($n = 6$) (**e**); IL-10 blockade (anti-IL-10 (5 μg ml^{-1}) and anti-IL-10R (5 μg ml^{-1}) antibodies) ($n = 6$) (**f**) or IL-6 blockade (tocilizumab, 50 μg ml^{-1}) ($n = 7$) (**g**). In **a** and **b**, the whiskers denote the minimum and maximum values, the box limits denote quartiles 1–3, and the centre bar denotes the mean. For **a**–**g**, significance was assessed using two-tailed unpaired *t*-tests; NS, not significant ($P > 0.05$).

STAT3 is a transcription factor that is activated downstream of multiple cytokines and growth factors[20]. When aberrantly phosphorylated, STAT3 contributes to lymphoproliferation, recurrent infections and increased

autoimmunity including eczema, type 1 diabetes and hypothyroidism[21]. Clinically, Mendelian *STAT3* gain of function (GOF) substantially overlaps with DS, indicating that engagement of STAT3 in CD4 T cells in DS may contribute to immune pathogenesis. The molecular mechanism of disease in STAT3 GOF is still debated, but STAT3 attenuation of STAT5 phosphorylation is the leading hypothesis. In DS, basal phosphorylated STAT5 (pSTAT5) levels were unaffected in CD4 T cells and even increased in CD4 TEMRAs (Extended Data Fig. 3b), suggesting that other factors are involved.

To test whether basal pSTAT3 is caused by cytokine signalling instead of an intrinsic STAT3 GOF in DS, we incubated whole blood with the FDA-approved Jak inhibitor tofacitinib. This treatment restored basal pSTAT3 to control levels, indicating that STAT3 activation in DS is Jak dependent (Fig. 2d and Extended Data Fig. 3c).

Given the detection of IFNα2 and IL-10 in a subset of individuals with DS (Fig. 1a), combined with the triplication of their cognate receptors (*IFNAR1*, *IFNAR2* and *IL10RB*) in DS, we hypothesized that basal pSTAT3 was due to heightened IFN-I or IL-10 response. However, IFN-I blockade did not affect basal pSTAT3 (Fig. 2e). Blocking IL-10 resulted in a modest decrease in pSTAT3 that was most pronounced in effector memory CD4 cells (Fig. 2f), suggesting that IL-10 may contribute to baseline signalling in DS. These blocking experiments were performed ex vivo, so the effect of these cytokines over time may not have been captured.

Having established that basal pSTAT3 is initiated upstream of Jaks and largely independent of IFN-I and IL-10, we turned to the IL-6–STAT3 axis. IL-6 is a potent inducer of pSTAT3 and was significantly elevated in DS (Fig. 1a,c). IL-6 blockade with the FDA-approved IL-6 receptor (IL-6R) inhibitor tocilizumab fully abrogated pSTAT3 (Fig. 2g and Extended Data Fig. 3d), indicating that IL-6 is a major mediator of baseline CD4 activation in DS. Furthermore, there was no hyper-response to exogenous IL-6 stimulation and no increase in IL-6R expression in DS (Extended Data Fig. 3e,f), confirming that elevated pSTAT3 is a result of high IL-6 rather than an intrinsic gain of expression or function in IL-6R. Together, these data implicate IL-6 signalling in ongoing CD4 T cell activation in DS and offer a

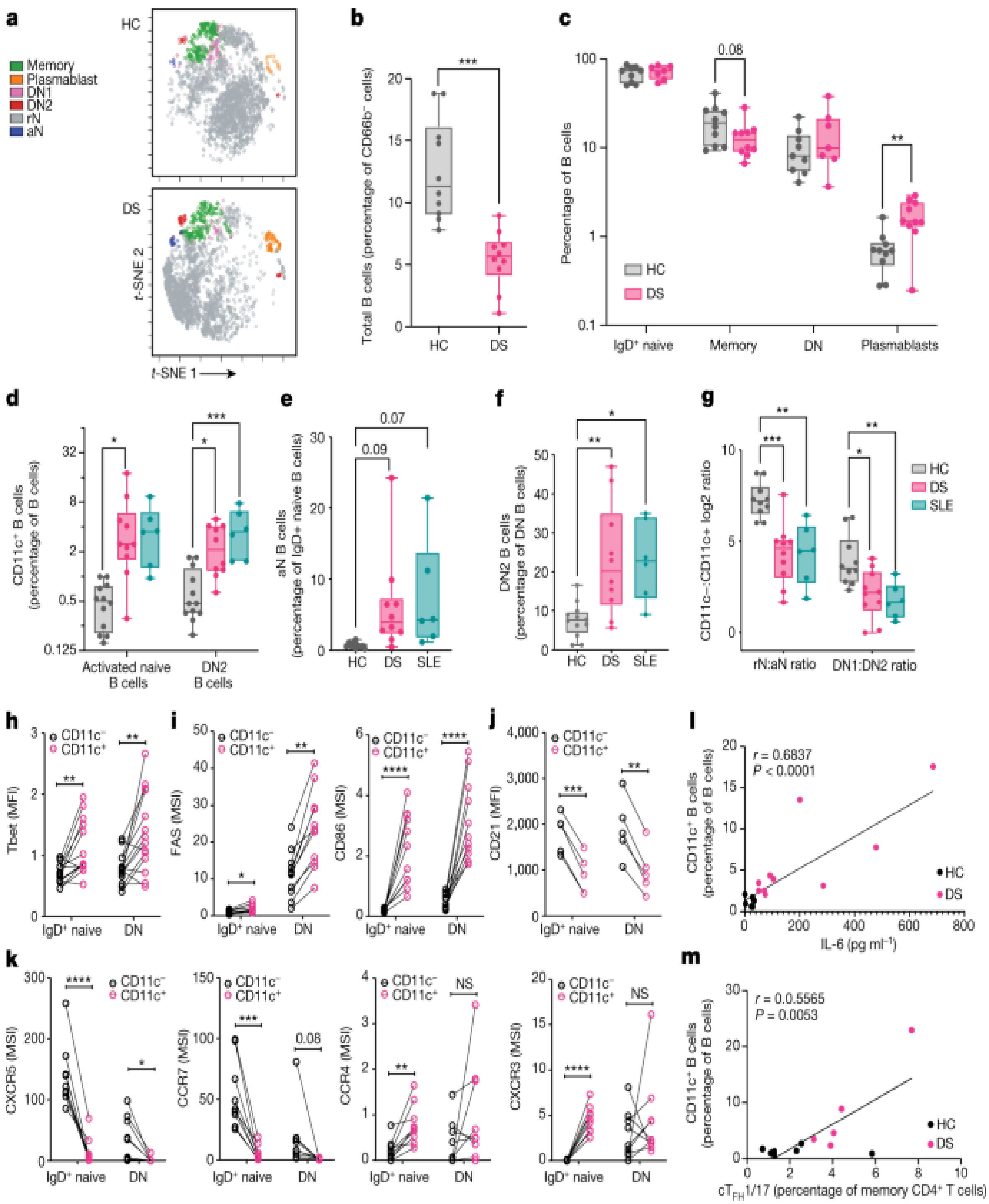
more specific therapeutic target than Jak inhibition for consideration by physicians treating individuals with DS.

Notably, genetic alterations of STAT3 signalling also lead to a disturbance of B cell subsets, with a decrease in a recently described atypical B cell activation in adults with *STAT3* loss of function and a corresponding increase in these $CD11c^+$ B cells in *STAT3* GOF. Given our findings of basal pSTAT3 in DS T cells, the overlap in clinical manifestations between DS and STAT3 GOF, and previously documented B cell disturbances in DS[13], we further investigated the B cell compartment.

Increased $CD11c^+Tbet^{high}CD21^{low}$ B cells

Total B cells were profoundly decreased in individuals with DS (Fig. 3a,b and Extended Data Fig. 4a) as seen previously[22,23]. All whole-blood samples were fixed before freezing, which enabled us to analyse all B cell populations including plasmablasts. The relative frequencies of B cell subsets were altered in DS: memory B cells were slightly decreased compared with HC individuals, whereas plasmablasts increased by almost threefold (Fig. 3a,c). Given the unexpected co-occurrence of elevated plasmablasts and depleted memory cells, combined with the simultaneous widespread elevation of pro-inflammatory cytokines and overactivation of CD4 T cells discussed above, we examined the possibility that B cells undergo inordinate activation in DS.

Fig. 3: Increased frequency of $CD11c^+Tbet^{high}CD21^{low}$ B cells in DS.



a, Representative *t*-distributed stochastic neighbour embedding (*t*-SNE) analysis of a fixed number of B cells to illustrate subset distribution in whole blood from adults with DS and age-matched HC individuals. **b**,**c**, The frequency in adults with DS ($n = 10$) and age-matched HC individuals ($n = 10$) of total B cells expressed as the percentage of $CD66b^{-}$ cells (non-

granulocytes) (**b**) and B cell subsets expressed as the percentage of total B cells (**c**). **d**–**g**, The frequency in HC adults ($n = 10$), adults with DS ($n = 10$) and patients with SLE ($n = 6$) of CD11c^{+} B cells in the IgD^{+} naive (CD27^{-}CD38^{low}IgD^{+}) or DN (CD27^{-}CD38^{low}IgD^{-}) compartments expressed as the percentage of total (**d**), IgD^{+} naive (**e**) or DN (**f**) B cells, and the log$_2$-transformed ratios of CD11c^{+} subsets to CD11c^{-} subset (rN:aN and DN2:DN1) (**g**). **h**–**k**, Intracellular Tbet expression (**h**), and surface expression of FAS and CD86 (**i**), CD21 (**j**) and CXCR5, CCR7, CCR4 and CXCR3 (**k**) in naive and DN B cells from both HC individuals ($n = 2$–4) and individuals with DS ($n = 3$–10). MSI, mean signal intensity; MFI, mean fluorescence intensity. **l**,**m**, Correlation of CD11c^{+} B cells and circulating IL-6 (**l**) and cT$_{FH}$1/17 (**m**). *r*, Pearson correlation coefficient. In **b**–**g**, the whiskers denote the minimum and maximum values, the box limits denote quartile 1 to quartile 3, and the centre bar denotes the mean. Significance was assessed using two-tailed unpaired *t*-tests (**b** and **h**–**k**) and one-way ANOVA with Tukey's post hoc analysis (**d**–**g**).

CD11c^{+} B cells are thought to derive from naive B cells by cytokine and T cell stimulation and/or TLR engagement outside of germinal centres[24]. They are also characterized by high Tbet and low CD21 expression and can differentiate into plasmablasts[25]. The two major documented subsets of CD11c^{+} B cells are CD27^{-}IgD^{+} activated naive (aN) B cells and a subpopulation of class-switched CD27^{-}IgD^{-} (double-negative (DN2)) B cells[26,27] (Extended Data Fig. 4b). This B cell activation is a hallmark of the heterogenous condition systemic lupus erythematosus (SLE). It has also been described in other autoimmune and autoinflammatory diseases including rheumatoid arthritis, ulcerative colitis and common variable immunodeficiency[25,27,28].

Our analysis revealed increased frequencies of CD11c^{+} B cells in both the IgD^{+} naive and double-negative (CD27^{-}IgD^{-}) compartments in individuals with DS compared with in HC individuals (Fig. 3d–f), at a frequency comparable to mild SLE (Fig. 3d–f). Elevation of previously described CD11c^{+} intermediates (aN and DN2 B cells) was concurrent with a reduction in traditional B cell activation, evidenced by decreased frequencies in resting naive (rN) and DN1 B cells. The rN:aN and DN1:DN2 ratios were

significantly lower in individuals with DS and SLE compared with in HC individuals (Fig. 3g). Despite the inherently variable cell frequencies in childhood, aN B cells were significantly higher in children with DS, and DN2 B cells trended upward (Extended Data Fig. 4c–g), demonstrating that this atypical B cell activation also occurs early in life.

We next performed high-dimensional analysis of functional markers of $CD11c^+$ B cells. As reported previously, both aN and DN2 B cells displayed upregulation of the IFNγ-induced transcription factor Tbet (Fig. 3h), and the surface receptors FAS (also known as CD95) and CD86 (also known as B7-2)[29] (Fig. 3i). These cells also exhibited lower expression of CD21 (Fig. 3j). Expression of chemokine receptors in $CD11c^+$ cells was consistent with previously described extrafollicular activation. The follicle-homing receptors CXCR5 and CCR7 were significantly downregulated in aN and DN2 B cells compared with in rN and DN1 B cells (Fig. 3k), concurrently with an increase in CXCR3 in aN cells (Fig. 3k), which suggests homing potential to inflamed tissues[30]. We also found elevated expression of a marker in aN B cells, CCR4 (Fig. 3k), a receptor that is classically associated with CD4 T_H2 cells and is thought to be involved in homing to the skin and lungs[31,32]. Further research is needed to elucidate whether CCR4 expression in these B cells can drive pathology.

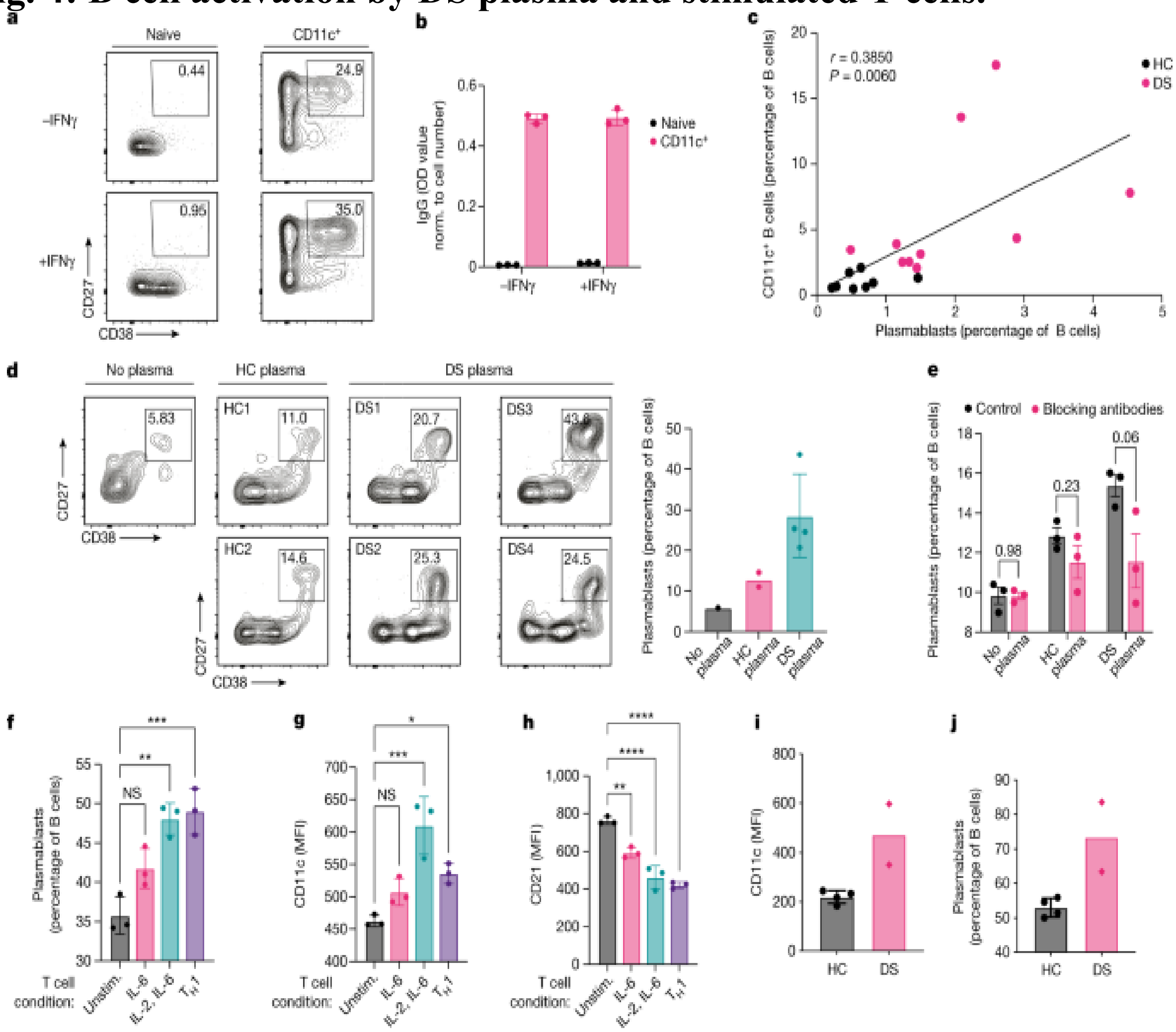
Although $CD11c^+$ B cell frequency in DS was not correlated with age[27,33] (Extended Data Fig. 4h), it was correlated with total cytokine levels (Extended Data Fig. 4i), specifically IL-6 levels (Fig. 3l), as well with the frequency of circulating T follicular helper 1/17 ($cT_{FH}1/17$) cells (Fig. 3m), as in other primary immunodeficiencies[28]. This suggests that cytokines and activated CD4 T cells contribute to putative extrafollicular B cell differentiation[24]. Finally, $CD11c^+$ B cell frequency was correlated with the number of autoimmune manifestations in our cohort (Extended Data Fig. 4j), indicating their potential role in autoimmunity.

In conclusion, $CD11c^+$ B cells are a prominent component of the abnormal B cell response in DS. On the basis of our research and previous work in other inflammatory conditions[25,28], we hypothesize that this dysregulated B cell response weakens immune tolerance and, ultimately, contributes to autoimmunity in DS.

Ex vivo naive B cell differentiation

CD11c$^+$ B cells can differentiate into antibody-secreting plasmablasts with little to no affinity maturation, potentially explaining their association with autoimmunity[34]. In vitro, we confirmed that stimulated CD11c$^+$ B cells sorted from HC peripheral blood mononuclear cells (PBMCs) differentiated into plasmablasts and secreted immunoglobulin G (IgG), in contrast to treated naive cells (Fig. 4a,b). This was potentiated by the addition of IFNγ (Fig. 4a), as previously published[27]. In whole blood, CD11c$^+$ B cell frequency was positively correlated with that of plasmablasts (Fig. 4c). We also confirmed that total IgG is elevated in DS plasma compared to HC plasma[13] (Extended Data Fig. 4k).

Fig. 4: B cell activation by DS plasma and stimulated T cells.



a,**b**, Plasma cell differentiation (**a**) and secreted IgG in the supernatant (**b**) after culture for 4 days of sorted HC naive or $CD11c^+$ B cells in the presence of BAFF, IL-2, IL-10, IL-21, the TLR7/8 ligand R848 with or without IFNγ. *n* = 3 biologically independent samples. Norm., normalized. **c**, Correlation of $CD11c^+$ B cell and plasmablast frequencies in adults DS (*n* = 12) and age-matched HC individuals (*n* = 8). **d**, Plasma cell differentiation after culture for 6 days of sorted HC naive B cells in the presence of BAFF, IL-2, IL-10, IL-21, R848 and IgG-depleted plasma from HC individuals (*n* = 2) or individuals with DS (*n* = 4). **e**, Magnetic-activated cell sorting (MACS)-isolated naive B cells from a healthy donor were cultured for 3 days with BAFF, IL-2, IL-21, R848, anti-IgM and plasma of HC individuals or individuals with DS in the presence of a combination of antibodies blocking IFN-I, IFN-II, IL-6 and TFNα signalling, in triplicates. Significance was assessed using two-tailed unpaired *t*-tests. Data are mean ± s.e.m. The results are representative of two independent experiments with *n* = 2 donors per group. **f**–**h**, Co-cultures containing T cells activated with IL-6, IL-2, both or polarized into T_H1 cells with IL-2, IL-12 and anti-IL-4 together with MACS-isolated naive B cells from the same donor, run in triplicates. The frequency of plasmablasts (**f**) and extracellular CD11c induction (**g**) and downregulation of CD21 expression (**h**) in non-plasmablast B cells after co-culture for 3–6 days. Significance was assessed using one-way ANOVA with Tukey's post hoc analysis. **i**,**j**, CD11c induction (**i**) and plasmablast differentiation (**j**) in naive $CD11c^-$ B cells isolated from controls (*n* = 4) or individuals with DS (*n* = 2) after 3 days of co-culture with CD4 T cells isolated from the same donors (syngeneic cultures). For **b**, **d** and **f**–**j**, data are mean ± s.d.

We next performed experiments to test the mechanisms underlying increased B cell activation in DS. To first address whether the cytokines in DS plasma could induce this differentiation, we co-incubated naive or total B cells sorted from control PBMCs in extrafollicular-stimulating conditions[27] in the presence of IgG-depleted plasma derived from HC individuals or individuals with DS. There was significantly more plasmablast differentiation in the presence of DS plasma compared to HC plasma (Fig. 4d and Extended Data Fig. 5a). Exogenous cytokine treatment could also influence plasmablast differentiation: IFNα and IFNγ potentiated differentiation, whereas IL-4 restricted it[35,36] (Extended Data Fig. 5b–d). Individual blocking of IL-6,

IFN-I, IFNγ and TNF in donor plasma had no effect, whereas Jak inhibitors inhibited plasmablast differentiation, suggesting that a few cytokines are involved (Extended Data Fig. 5f). Blocking a combination of four cytokines (IL-6, IFN-I, IFNγ and TNF) reversed the DS plasma-induced augmentation of plasmablast differentiation (Fig. 4e). Thus, the cytokines present in DS plasma—at least, IL-6, IFN-I, IFNγ and TNF—are drivers of naive B cell differentiation into plasmablasts.

Given the mounting evidence in vivo that T cells drive extrafollicular B cell activation[24,28], together with our findings of basal IL-6 signalling in DS CD4 T cells, we tested whether T cells contribute to this B cell response. Previous studies have demonstrated that T_H1-polarized T cells can induce extrafollicular differentiation of naive B cells in vitro[33,37]. When we modelled DS CD4 T cell activation with exogenous IL-6, co-culture of naive B cells and T cells pretreated with IL-2 and IL-6 resulted in plasmablast differentiation equivalent to that of T_H1–B cell co-cultures (Fig. 4f). Exogenous IL-6 alone did not affect plasmablast differentiation (Extended Data Fig. 5b). Furthermore, T_H1-cell- and IL-6-primed CD4 T cells induced CD11c expression in co-cultured B cells (Fig. 4g), together with CD21 downregulation (Fig. 4h) and a slight increase in Tbet (Extended Data Fig. 5g). Although IFNγ is thought to have a major role in the T_H1-mediated atypical B cell differentiation and was indeed present in the T_H1 supernatants, we did not detect IFNγ in the IL-6 conditions (Extended Data Fig. 5h), which indicates that other cytokines drive B cell activation in these conditions. Together, these results demonstrate that cytokines and T cells in combination can drive an extrafollicular B cell response.

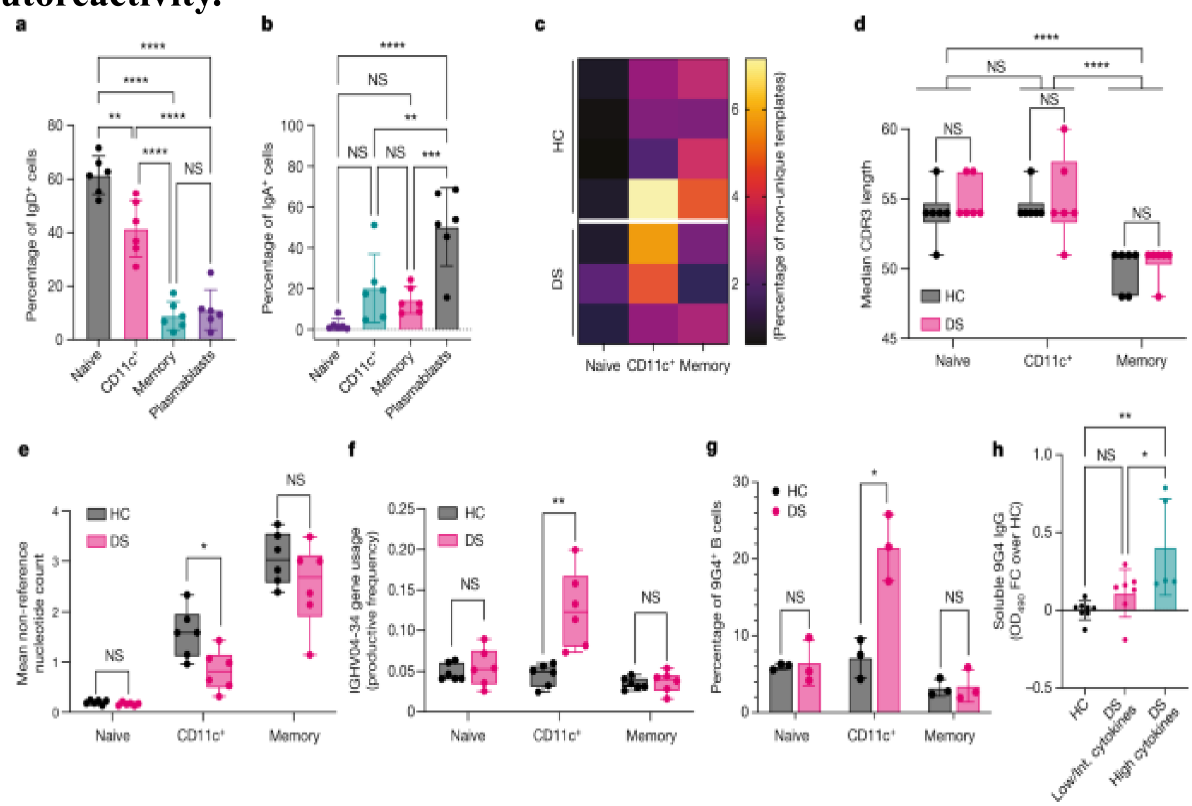
To better replicate the physiological conditions of DS, we performed these experiments with pre-incubation of T cells in plasma derived from individuals with DS. We found that these cells induced increased plasmablast differentiation of naive B cells compared with T cells incubated in control plasma (Extended Data Fig. 5i). Finally, to determine whether activated CD4 T cells of individuals with DS are poised to induce differentiation of naive B cells, we isolated CD4 T cells from control individuals ($n = 4$) and individuals with DS ($n = 2$) and co-cultured them with $CD11c^-$ naive B cells in syngeneic co-cultures. Without exogenous polarization, T cells from individuals with DS induced more CD11c

expression and plasmablast differentiation than those of HC individuals (Fig. 4i,j). These data indicate that cytokine milieu and steady-state cellular activation contribute to atypical B cell activation in DS.

Autoimmune features in CD11c^{+} B cells

Next, we looked at immunoglobulin isotype expression to further ascertain the antibody-secreting potential and the naivety of CD11c^{+} B cells. The frequency of IgD^{+}CD11c^{+} B cells was intermediate between naive and memory B cells (Fig. 5a). The proportion of IgA^{+}CD11c^{+} B cells was similar to memory B cells (Fig. 5b), demonstrating that a portion of these atypical cells have gone through class switching and are therefore probably antigen experienced. We did not detect significant differences in isotype usage between HC and DS CD11c^{+} B cells (Extended Data Fig. 6a).

Fig. 5: CD11c^{+} B cells from individuals with DS are more prone to autoreactivity.



a,**b**, Expression of IgD (**a**) and IgA (**b**) in naive, $CD11c^+$ and memory B cells and plasmablasts from adults with DS and age-matched HC individuals. $n = 3$ each. **c**–**h**, BCR sequencing analysis of genomic DNA isolated from sorted naive, $CD11c^+$ and memory B cells from controls ($n = 6$) and individuals with DS ($n = 6$). **c**, The fraction of productive BCRs represented more than once in each sample. Individuals from whom a sample had fewer than 1,000 productive templates were excluded. **d**, CDR3 length of productive BCRs in B cells subsets in HC and DS, expressed as the number of nucleotides (nt). **e**, The mean number of nucleotides different from reference in the V gene of productive BCRs in B cells subsets in HC and DS. **f**, The frequency of productive BCRs that were aligned to the *IGHV4-34* gene in each sample. **g**, 9G4 surface expression in naive, $CD11c^+$ and memory B cells from HC individuals ($n = 3$) and individuals with DS ($n = 3$). **h**, ELISA quantification of 9G4 antibodies in the plasma of HC individuals ($n = 8$), and individuals with DS in the low/medium ($n = 7$) and high ($n = 5$) cytokine groups, expressed as the fold change over HC individuals. OD_{490}, optical density at 490 nm. For **a**, **b**, **g** and **h**, data are mean ± s.d. For **d**–**f**, the whiskers denote the minimum and maximum values, the box limits denote quartile 1 to quartile 3, and the centre bar denotes the mean. Significance in **a** and **h** and significance between cell subsets in **d** was assessed using one-way ANOVA with Tukey's post-hoc analysis. Significance in **d**–**g** between the HC and DS groups was assessed using two-tailed paired *t*-tests.

To assess the clonality of $CD11c^+$ B cells, we performed B cell receptor (BCR) sequencing (BCR-seq) of DNA from sorted naive, $CD11c^+$ and memory B cells from individuals with DS ($n = 6$) and age-matched control individuals ($n = 6$) at steady state (Extended Data Fig. 6b,c). In naive cells, around 99% of BCRs were unique, whereas, in $CD11c^+$ cells and memory cells, up to 7% of BCRs were expanded (Fig. 5c), suggesting that, like memory B cells, $CD11c^+$ B cells undergo clonal expansion (or are the result of clonal expansion) rather than non-specific stimulation. There was no difference in clonality in $CD11c^+$ B cells between HC individuals and individuals with DS, as evidenced by similar fractions of non-unique BCR templates (Fig. 5c) and similar Simpson diversity index metrics (Extended Data Fig. 6d). The complementarity determining region 3 (CDR3) repertoires of $CD11c^+$ and memory B cell subsets overlapped at the

nucleotide and amino acid levels (Extended Data Fig. 6e,f), demonstrating that cells in these two phenotypically defined subsets derive from the same lineage. In future studies, BCR-seq analysis of donors immunized with a known antigen will enable closer examination of clonal expansions of these B cells.

The median CDR3 length of $CD11c^+$ B cells was similar to that of naive cells and significantly longer than that of memory cells (Fig. 5d). It was not significantly different between these cells in individuals with DS and HC individuals. Increased CDR3 length in antibody-secreting cells is associated with antibody polyreactivity and autoimmunity[38], suggesting that these $CD11c^+$ B cells are prone to autoreactivity.

The number of non-reference nucleotides in the V genes of $CD11c^+$ B cells was intermediate between that of naive and memory B cells (Fig. 5e). Notably, $CD11c^+$ B cells in individuals with DS had significantly fewer non-reference nucleotides compared with in HC individuals. This may indicate a comparative lack of somatic hypermutation, which could lead to a broader, more non-specific humoral response. Our method was limited to CDR3 sequencing; thus, characterization of the full antigen-binding region is still needed to ascertain the true somatic hypermutation rate.

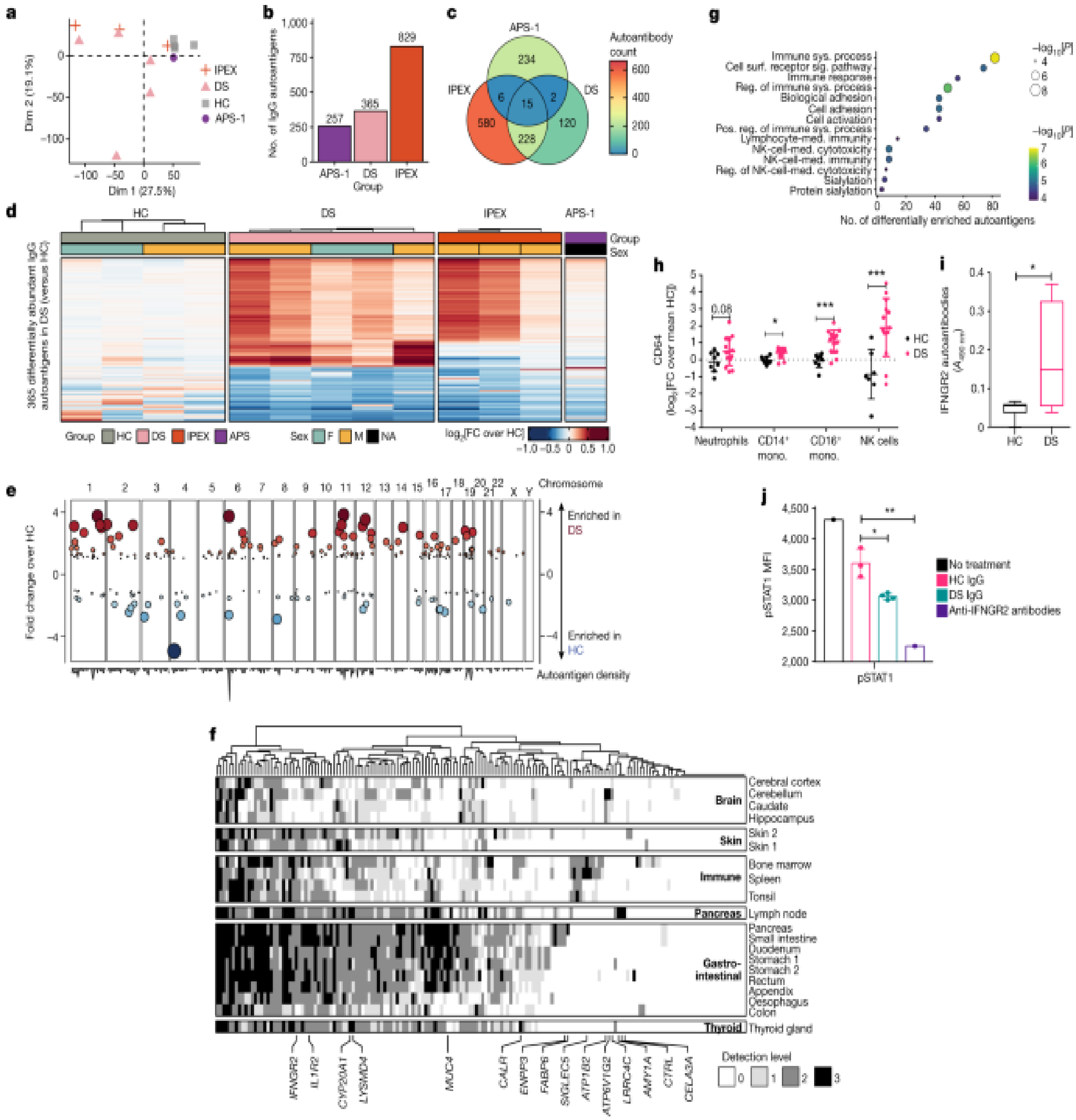
Analysis of BCR V gene usage (Extended Data Fig. 6g) revealed that $CD11c^+$ B cell expansion is probably more autoreactive in individuals with DS than in HC individuals. There was a significantly higher usage of the *IGHV4-34* gene, which is associated with autoreactivity[27,38], in $CD11c^+$ B cells in DS (Fig. 5f). We confirmed these BCR-seq results using flow cytometry: 9G4 idiotype antibodies encoded by this *IGHV4-34* gene segment were more highly expressed in these atypical B cells in DS (Fig. 5g). In vitro stimulation of sorted $CD11c^+$ B cells into antibody-secreting cells led to higher 9G4 secretion than from memory B cells (Extended Data Fig. 4h), further accentuating the link between these rare B cells and autoimmune potential. Circulating 9G4 antibodies were also elevated in the plasma of individuals with DS and were more abundant in the high-cytokine individuals with DS than in the low/intermediate groups (Fig. 5h). 9G4 antibodies are known contributors of autoimmunity in SLE, displaying specificity for nuclear antigens, dsDNA and apoptotic cells[39,40]. In

conclusion, $CD11c^+$ B cells in DS are present at a higher frequency and are more likely to exhibit features of self-reactive BCRs.

Autoantibodies are enriched in DS

Given clinical autoimmunity in DS and our findings of cytokine dysregulation and autoimmune-prone B cells, we hypothesized that trisomy 21 results in the generation of autoantibodies. To characterize the DS autoreactive repertoire, we assessed the plasma IgG and IgA reactivity of individuals with DS ($n = 5$), age-matched healthy individuals ($n = 4$), 3 individuals with immunodysregulation polyendocrinopathy enteropathy X-linked syndrome (IPEX) and 1 individual with autoimmune polyglandular syndrome type 1 (APS-1) (disease controls; $n = 4$) against >21,000 conformationally intact human proteins (CDI HuProt protein microarray) (Fig. 6a and Extended Data Fig. 7a–d).

Fig. 6: Autoantibody repertoire in individuals with DS.



a, PCA analysis of the HuProt IgG dataset for adults with DS ($n = 5$), age-matched HC individuals ($n = 4$), and patients with IPEX ($n = 3$) and APS-1 ($n = 1$). **b**, The number of IgG autoantigens enriched at least twofold in individuals with DS, IPEX and APS-1 compared with HC individuals. **c**, Enriched IgG autoantigens overlapping between disease groups. **d**, Enriched IgG autoantigens in HC individuals and in individuals with DS, IPEX and APS-1. The colour intensity corresponds to the $\log_2$-transformed fold change expression value relative to the mean of healthy adult controls. F, female; M, male; NA, unknown. **e**, Chromosomal expression pattern of IgG autoantigens enriched in DS. **f**, Gene expression pattern of IgG autoantigens

enriched in DS ($n = 5$) according to the Human Protein Atlas. **g**, GO analysis of IgG autoantigens enriched in DS ranked by the number of autoantigens found to be enriched in the associated gene set. The dot size and colour intensity correspond to the FDR-adjusted *P* value. Med., mediated; NK, natural killer; reg., regulation; sig, signalling; surf, surface; sys., system. **h**, The surface expression of CD64 in monocytes, natural killer cells from individuals with DS ($n = 14$) and age-matched HC individuals ($n = 8$). Mono, monocytes. **i**, ELISA analysis of anti-IFNGR2 autoantibodies in the plasma from adults with DS ($n = 4$) and age-matched HC individuals ($n = 7$). **j**, Neutralizing IFNγ signalling in THP-1 cells by IgG fraction of plasma from adults with DS ($n = 3$), age-matched HC individuals ($n = 3$) or recombinant anti-IFNGR2 antibody. Significance was assessed using one-way ANOVA with Tukey's post hoc analysis. For **h** and **i**, significance was assessed using two-tailed unpaired *t*-tests. For **h** and **j**, data are mean ± s.d. For **i**, the whiskers denote the minimum and maximum values, the box limits denote quartile 1 to quartile 3, and the centre bar denotes the mean.

Principal component analysis (PCA) grouped DS samples together and away from HC individuals (Fig. 6a), indicating differential self-antigen binding. The DS samples clustered with samples from patients with IPEX ($n = 3$), and away from the sample from the individual with APS-1 ($n = 1$) (caused by mutation in *AIRE*), although the low sample number is an important caveat (Fig. 6a).

Analysis of differentially abundant autoantigens in individuals with DS compared with HC individuals ($\log_2$-transformed fold change > 1, $P < 0.05$) yielded 365 proteins, in contrast to 257 and 829 proteins for APS-1 and IPEX, respectively (Fig. 6b). This indicates that autoimmunity in trisomy 21 is on par with bona fide autoimmune diseases. We observed little overlap of autoantigens between DS and APS-1 (2 out of 257), whereas a large proportion of the DS autoantigens overlapped with those in IPEX (228 out of 829) (Fig. 6c,d), a syndrome of severe multiorgan inflammation caused by defects in T regulatory (T_{reg}) cells. The highly similar autoantibody repertoires points to T cell overactivation as a possible driver of autoreactive B cells in DS. Consistent with this finding, absolute T_{reg} cell counts were decreased in individuals with DS compared with in HC individuals, despite similar T_{reg} cell frequency[41] (Extended Data Fig. 2f,g). Finally, 120

autoantigens were unique to DS, including metabolic (TPH1, ATP5F1, PLPP2), cell signalling (RRAGC, GPR143, PTGER4), immune signalling (DTX4, CCL11, TROVE2) and neuronal pathway intermediates (ABAT, ATP6V1G2, KCTD7) (Fig. 6c,d and Supplementary Data 1 and 2).

The autoantigens enriched in DS were encoded across all 23 chromosomes (Fig. 6e). They included proteins expressed at sites of clinical autoimmunity in DS, such as the gastrointestinal tract (MUC4, ENPP3, FABP6) and the pancreas (AMY1A, CELA3A, CTRL) (Fig. 6f). The four individuals with DS with hypothyroidism assayed did not have known anti-thyroid autoantibodies in contrast to previously described individuals with DS[42] but, instead, had antibodies directed against proteins non-specifically expressed in the thyroid (CALR, LYSMD4, CYP20A1) (Fig. 6f). We also saw an enrichment of autoantigens expressed in the central nervous system (CNS) (ATP1B2, LRRC4C, ATP6V1G2) (Fig. 6f). Although antibodies cannot cross the intact blood–brain barrier, pathology can be mediated through a weakened blood–brain barrier[43]. Erroneous expression of normally neuronally restricted markers through altered gene expression in DS is another potential mechanism[44]. Given the prominent central nervous system features of DS, further investigation is warranted. Finally, immune autoantigens were prominently enriched in DS, including IFNGR2, TLR9, CD1C and IL-1R2 (Fig. 6f,g and Extended Data Fig. 7e). Immune system process, cell-surface receptor signalling and immune response were the most significantly enriched Gene Ontology (GO) terms (Fig. 6g). Abundant autoantibodies directed against the immune system may promote further immune dysregulation.

On the cellular side, we detected increased expression of high-affinity FCGR1A (also known as CD64) in monocytes and natural killer cells in DS (Fig. 6h). This receptor can engage bound antibodies and immune complexes to trigger potent inflammation and tissue injury[45]. Cells in individuals with DS may be poised to cause tissue damage based on the observed expression patterns.

We next validated the HuProt array results using an enzyme-linked immunosorbent assay (ELISA), confirming the presence of anti-IFNGR2, MSTN and ATP6V1G2 autoantibodies in individuals with DS but not in HC

individuals (Fig. 6i and Extended Data Fig. 7f). We then functionally tested anti-IFNGR2 autoantibodies (enriched 2.7-fold in DS, false-discovery rate (FDR) = 0.01): compared to HC IgG, DS IgG resulted in a significant decrease in the response to IFNγ as measured by pSTAT1 (Fig. 6j). This demonstrates that immune-targeting autoantibodies in individuals with DS can directly inhibit the activity of cytokines. As genetic defects in IFNγ signalling cause susceptibility to mycobacterial disease[46], anti-IFNGR2 autoantibodies may predispose individuals with DS to mycobacterial disease, which should be investigated in DS especially as individuals age.

Given the increased susceptibility to severe COVID-19 in individuals with DS[47], we examined anti-IFN-I autoantibodies. These were modestly enriched in individuals with DS compared with HC individuals in the CDI array, but much less so than in APS-1 (Extended Data Fig. 7g). We confirmed anti-IFNα2 and anti-IFNω autoantibody enrichment in DS by Gyros assay, at titres over tenfold lower than that of APS-1 samples (Extended Data Fig. 7h). Plasma from individuals with DS with or without COVID-19 was insufficient to inhibit IFN-I, whereas samples from the individual with APS-1 fully abrogated IFN-I signalling. We noted only mild neutralization of very low amounts of IFNα2 (100 pg ml^{-1}) by DS plasma (Extended Data Fig. 7i).

In conclusion, we detected a prominent autoantibody landscape in the plasma of individuals with DS. Together with our findings of B and T cell activation, we hypothesize that this broad humoral response predisposes individuals with DS to form high-affinity, tissue-specific antibodies that manifest clinically as autoimmune disease.

Discussion

Autoantibodies in DS have long been implicated in many of the syndrome's features, such as thyroid disease, type 1 diabetes and even cognitive decline[42,48,49,50]. Here we delineate the predisposition to autoreactivity and identify the autoantibody repertoire in DS (Extended Data Fig. 8). Two of the dysregulated immune features that we detected—the cytokine landscape and $CD11c^+$ B cell frequency—had a direct moderate association with clinical immune manifestations ($P = 0.0141$ by ANOVA and $P = 0.0795$ by

r^2 test, respectively). Perhaps dampening chronic steady-state inflammation in DS could prevent the accumulation of self-reactive autoantibodies and lessen disease burden.

We identified an increased frequency of CD11c$^+$Tbet$^+$CD21^{low} B cells as a hallmark of immune dysregulation in DS and established a link between these cells and the elevated plasmablast frequency and increased IgG in DS plasma. Others have also implicated these cells in the production of self-reactive antibodies[34], which may partly explain autoimmune susceptibility in DS. Beyond the increased frequency of these intrinsically pro-inflammatory cells in DS, BCR sequencing revealed that CD11c$^+$ cells are more likely to have self-reactive features. Thus, in DS, these putatively extrafollicular B cells are skewed towards autoimmunity in both quantity and quality. Their direct involvement in the generation of autoantibodies and their potential damage at the tissue level merit further investigation.

We uncovered activation of naive B cells not only by cytokines in DS plasma but also by CD4 T cells with basal pSTAT3. Patients with *STAT3* loss-of-function mutations have fewer CD11c$^+$ B cells, whereas those with *STAT3* GOF mutations have more[28], like individuals with DS. Although it is unclear whether the exaggerated naive B cell activation in *STAT3* GOF is due to intrinsic STAT3-mediated signalling in B cells, stimulation by overactive T cells or both, we hypothesize that baseline T cell activation in DS has an important role in this B cell response. This notion is further emphasized by the similarity in autoantibody repertoires between DS and IPEX, an autoimmune disease mediated by lack of T cell regulation. Our in vitro experiments demonstrate that IL-6 activation of T cells is sufficient to drive the differentiation of naive B cells into CD11c$^+$ B cells and plasmablasts. Although the physiological implications of this model must be firmly tested, these data provide a mechanistic link between IL-6 activation of T cells and this atypical B cell response. Thus, resolving basal T cell activation in DS with a Jak inhibitor or IL-6 blockade is a promising therapeutic avenue to temper hyperinflammation and the autoimmune feed-forward loop that we describe.

Methods

Sample collection

The study was approved under protocols at Mount Sinai Health System (MSHS) (IRB-18-00638/ STUDY-18-00627 and IRB-20-03276), Boston Children's Hospital (04-09-113R), National Institute of Allergy and Infectious Disease (NIAID, NIH) (05-I-0213), Rockefeller University (JCA-0700 and XFK-0815), the French Ethics Committee Comité de Protection des Personnes, the French National Agency for Medicine and Health Product Safety and the Institut National de la Santé et de la Recherche Médicale (protocols C10-13 and C10-14). Patients were approached in person at the hospital or by email obtained through the NIH's DS-Connect national registry (https://dsconnect.nih.gov). All patient data were deidentified. Written informed consent for all individuals in this study was provided in compliance with an institutional review board protocol. All samples from uninfected patients were drawn in the context of an outpatient routine visit and patients exhibited no signs of infection (fever, runny nose, cough, sore throat). Demographics and characteristics of individuals with DS are outlined in Extended Data Tables 1 and 2. For the six patients with SLE included in Fig. 3 and Extended Data Fig. 4, four had severe disease with renal involvement, one had moderate disease without renal involvement and one had mild disease with renal involvement that was resolved at time of sampling (Extended Data Table 3). Half of the patients had active disease and the other three donors were in remission on treatment. From each donor, blood was drawn into a cell preparation tube with sodium heparin (BD Vacutainer). Plasma was isolated from Ficoll separation and stored at −80 °C until use. Whole blood was either directly fixed using Proteomic Stabilizer PROT1 (SmartTube) and frozen at −80 °C, or fresh whole-blood samples were stained by adding 270 µl blood directly to a MDIPA tube containing the lyophilized antibody panel, mixing and incubating for 30 min at room temperature, followed by stabilization with PROT1 and freezing.

Determination of clinical immune dysfunction score

A clinical questionnaire was filled out by patients' legal guardian or referring physician. One point was administered for each autoimmune manifestation or infectious episode. In cases in which an episode (such as pneumonia) was characterized as recurrent or reported more than three

times, we capped the score for this category at three points. For example, if a patient had hypothyroidism and 12 ear infections, the total immune clinical score for this patient would be 4 points.

Mass cytometry

The detailed protocol for mass cytometry was previously described[51]. Frozen stabilized blood samples were thawed according to the manufacturer's recommended protocol, then washed with barcode permeabilization buffer (Fluidigm). The samples were uniquely barcoded using the Cell-ID 20-Plex Pd Barcoding Kit (Fluidigm) and pooled together. For previously unstained samples, cells were incubated with an antibody cocktail for surface markers to identify major immune populations, followed by methanol permeabilization, heparin-block and stain with a cocktail of antibodies against intracellular targets, including markers of phosphorylation and signalling. All of the following antibodies were purchased from Fluidigm: Cd111-granzyme B, Cd112-IgA, In113-CD57, In115-CD11c, Cd116-IgD, I127-127I, Ce140-140Ce, Pr141-Ki67, Nd142-CD19, Nd143-CD45RA, Nd144-CD103, Nd145-CD4, Nd146-CD8, Sm147-pSTAT5, 150Nd-pSTAT5, Nd148-CD16, Sm149-CD127, Sm149-pSTAT6, Nd150-CD1c, Eu151-CD123, Sm152-CD66b, Eu153-pSTAT1, Sm154-ICOS, Gd155-CD27, Gd156-p38, 158Gd-pSTAT3, Tb159-pMAPKAP2, Gd160-CD14, Dy161-CD56, Dy162-TCRgd, Dy162-CD169, Dy163-CD172a/b, Dy164-CD69, Ho165-CD64, Ho165-STAT3, Er166-CD25, Er167-pERK1/2, Er168-CD3, Tm169-CD71, Tm169-STAT1, Er170-CD38, Yb171-CD95, Yb171-CD141, Yb172-CD39, Yb173-Tbet, Yb174-HLADR, Lu175-pS6, Yb176-CD54, Pr141-IFNγ, Nd144-CD141, 171Yb-CD141, Sm147-IL-1B, Sm149-IL-1RA, Eu153-TNF, Gd156-IL-6, Gd158-IL-2, Tb159-GM-CSF, Dy164-IL-17A, Ho165-CCL4, Er166-IL-10, Tm169-IFNα2b, Yb173-IL-8, Lu175-IL-29 and Yb176-CXCL10. After washing, cells were incubated in freshly diluted 2.4% formaldehyde containing 125 nM Ir Intercalator (Fluidigm), 0.02% saponin and 30 nM OsO4 (ACROS Organics) for 30 min at room temperature. Samples were then washed and acquired immediately.

For acquisition, samples were washed with PBS + 0.2% BSA, PBS and then CAS buffer (Fluidigm). The final solution in CAS buffer consisted of 1 million cells per ml and a 1/20 dilution of EQ beads (Fluidigm). After

routine instrument optimization, samples were acquired at a rate of <300 events per second on a Helios mass cytometer (Fluidigm) with a modified wide-bore injector (Fluidigm).

FCS files of acquired events were normalized and concatenated with Fluidigm acquisition software, deconvoluted using a MATLAB-based debarcoding application and the resulting files were analysed using Cytobank. Cell events were identified as Ir191/193-positive and Ce140-negative events. Doublets were excluded on the basis of Mahalanobis distance and barcode separation and with the Gaussian parameters calculated using the Helios CyTOF software. Downstream data analysis was performed on Cytobank, by biaxial gating of immune populations according to a previously published gating scheme[51] outlined in Supplementary Fig. 1. All cell frequency analyses were restricted to adult HCs and adults with DS (aged at least 18 years) unless otherwise indicated.

Flow cytometry

B cell phenotyping

Cryopreserved PBMCs were thawed and allowed to rest briefly in complete RPMI medium supplemented with 10% FBS, 1% penicillin–streptomycin and 1% GlutaMax. Cells were immunostained with antibodies in 0.5% BSA in PBS for 1 h, washed 3 times in 0.5% BSA in PBS for 1 h and acquired immediately. The following antibodies were used: CD19 APC-Cy7 (SJ25C1), CD27 FITC (M-T271), CD38 APC (HIT2), CD38 PE-Cy7 (HIT2), CD11c PE (B LY6), IgD BV421 (IA6), CD21 APC (Bu32) and anti-human 9G4 IgG APC (provided by J. Farmer).

pSTAT1 staining

The IgG fraction was isolated from HC and DS plasma using the Protein G HP SpinTrap (Sigma-Aldrich). THP-1 cells were incubated with the IgG fraction diluted 1:3 in complete RPMI with 10% FBS for 20 min or recombinant anti-IFNGR2 antibodies (2 μg ml^{-1}, Thermo Fisher Scientific, PA5-47938) at 37 °C for 20 min followed by stimulation with IFNγ (10 ng ml^{-1}) for 20 min at 37 °C. Cells were stained with Zombie Aqua Live-Dead

(BioLegend) for 30 min on ice, then fixed/permeabilized in 90% ice-cold methanol and stained with anti-phospho-STAT1-PE (1:25, BD) for 1 h on ice. Flow cytometry was acquired on the BD LSR Fortessa II system, and data were analysed with FlowJo.

Whole-blood cytokine blocking experiments

Whole blood was incubated for 4 h with tofacitinib (500 nM, ApexBio) or tocilizumab (50 μg ml^{-1}, Selleckchem) or vehicle controls (DMSO and human IgG control, 50 μg ml^{-1}, BioLegend, respectively). For IFN-I blocking, whole blood was incubated with a combination of anti-IFNAR2 (2.5 μg ml^{-1} PBL Assay Science), anti-IFNα (0.2 μg ml^{-1}, PBL 31110–1) and IFNβ (0.2 μg ml^{-1}, PBL 31401-1) antibodies or vehicle control for 4 h. For IL-10 blocking, whole blood was incubated with a combination of anti-IL-10 (5 μg ml^{-1}, BioLegend) and anti-IL-10R (5 μg ml^{-1}, BioLegend) for 4 h. After blocking, whole blood was fixed using Proteomic Stabilizer PROT1 (1.4× blood volume, SmartTube) for 10 min at room temperature and frozen at −80 °C.

B cell stimulation

For plasmablast differentiation of $CD11c^+$ cells (Fig. 4a), 2–5 × 10^4 naive ($CD19^+CD38^-CD27^-CD11c^-$) and $CD11c^+$ ($CD19^+CD11c^+$) B cells were sort-purified from the PBMCs of healthy donors using the IMI5L sorter (BD) and rested overnight in complete RPMI. They were then stimulated with IL-21 (50 ng ml^{-1}; BioLegend), BAFF (100 ng ml^{-1}, BioLegend), IL-10 (250 ng ml^{-1}, BioLegend), IL-2 (5 ng ml^{-1}) and R848 (1 μg ml^{-1}; Invivogen) with or without IFNγ diluted in complete RPMI. Cells were collected at day 4 and stained for flow cytometry.

For the cytokine-induced plasmablast differentiation of B cells (Fig. 4d,e and Extended Data Fig. 5a–f), 2–5 × 10^4 sorted naive or total B cells were sorted from healthy donor PBMCs using fluorescence-activated cell sorting, the Pan B Cell Isolation Kit II or the Naive B Cell Isolation Kit II (Miltenyi Biotech), as indicated. They were cultured for 3–6 days in complete RPMI supplemented with IL-21 (50 ng ml^{-1}; BioLegend), BAFF (10 ng ml^{-1},

BioLegend), IL-2 (50 ng ml^{-1}, BioLegend), R848 (1 μg ml^{-1}; Invivogen) with or without IL-10 (250 ng ml^{-1}, BioLegend) and/or anti-human IgM (5 μg ml^{-1}, Southern Biotech) as indicated, in addition to either (1) plasma from HC individuals or individuals with DS that was depleted from IgG using the Protein G HP SpinTrap (Sigma-Aldrich) at a final dilution of 1:4 or (2) exogenous addition of IL-6 (100 ng ml^{-1}, BioLegend), IFNα2b (100 U ml^{-1}, Merck), IL-4 (20 ng ml^{-1}, BioLegend) and IFNγ (20 ng ml^{-1}, BioLegend). At least two healthy donors and four individuals with DS were tested for each experiment. Cells were collected at indicated days and stained for flow cytometry.

For cytokine-blocking experiments, sorted naive B cells were cultured as described above in the presence of the following blocking antibodies: anti-IFNGR2 (2 μg ml^{-1}, Thermo Fisher Scientific, PA5-47938), anti-IFNAR2 (2.5 μg ml^{-1} PBL Assay Science), anti-IFNα (0.2 μg ml^{-1}, PBL 31110–1) and IFNβ (0.2 μg ml^{-1}, PBL 31401-1), tocilizumab (50 μg ml^{-1}, Selleckchem), adalimumab (2 μg ml^{-1}, Selleckchem) or a combination. Cells were collected at indicated days and stained for flow cytometry.

T cell–B cell co-cultures

For co-cultures of in vitro polarized T cells (Fig. 4f–h and Extended Data Fig. 5g–i). CD4 T cells were purified using MACS (Miltenyi Biotec) from healthy donor PBMCs and cultured in the presence of plate-bound anti-CD3 (OKT3, 10 μg ml^{-1}) and anti-CD28 (CD28.2, 5 μg ml^{-1}) antibodies in complete RPMI alone (unstimulated control) or with (1) IL-2 (50 ng ml^{-1}), IL-12 (5 ng ml^{-1}) and anti-IL-4 (1 μg ml^{-1}, R&D MAB204-SP) (T_H1 conditions); (2) IL-6 (50 ng ml^{-1}) with or without IL-2 (50 ng ml^{-1}); or (3) plasma from HC individuals or individuals with DS (1:2 ratio). After 4 days, T cells were washed and irradiated (5,000 rad) and co-cultured with freshly MACS-sorted naive B cells from the same donor. Cells were co-cultured for 3–6 days in complete RPMI supplemented with IL-21 (50 ng ml^{-1}), IL-2 (50 ng ml^{-1}) and anti-human IgM (5 μg ml^{-1}, Southern Biotech) and collected for flow cytometry. The supernatants were collected for IFNγ ELISA quantification (BioLegend) according to the manufacturer's instructions.

For T cell–B cell co-cultures from HC donors and donors with DS (Fig. 4i,j), CD4 T cells were MACS-purified and co-incubated with sort-purified $CD19^{+}CD38^{-}CD27^{-}CD11c^{-}$ B cells on the same day. B and T cells were co-cultured for 3 days in complete RPMI supplemented with IL-21, IL-2 and anti-human IgM as above and collected for flow cytometry.

BCR analysis

Sample preparation

Donor PBMCs were thawed in complete RPMI and rested overnight. Naive ($CD19^{+}CD27^{-}CD38^{low}CD11c^{-}$), $CD11c^{+}$ ($CD19^{+}CD27^{-}CD38^{low}CD11c^{-}$) and memory ($CD19^{+}CD27^{+}CD38^{-}$) B cells were sort-purified using an IMI5L sorter (BD) and genomic DNA was isolated with QIAamp DNA Blood Mini Kit (Qiagen).

Library preparation

Sample data were generated using the immunoSEQ Assay (Adaptive Biotechnologies). The somatically rearranged *Homo sapiens* BCR IgH locus *CDR3* was amplified from genomic DNA using a two-step, amplification-bias-controlled multiplex PCR approach[52,53]. The first PCR consists of forward and reverse amplification primers specific for every V and J gene segment and amplifies the hypervariable complementarity-determining region 3 (*CDR3*) of the immune receptor locus. Furthermore, primers are included to amplify B cell precursor DJ rearrangements. The second PCR adds a proprietary barcode sequence and Illumina adapter sequences[54]. *CDR3* libraries were sequenced on the Illumina instrument according to the manufacturer's instructions.

Data analysis

Raw sequencing reads were demultiplexed according to Adaptive's proprietary barcode sequences. Demultiplexed reads were then processed to remove adapter and primer sequences, identify and remove primer dimer, germline and other contaminant sequences. The filtered data are clustered

using the relative frequency ratio between similar clones and a modified nearest-neighbour algorithm to merge closely related sequences to correct technical errors introduced through PCR and sequencing while allowing for somatic hypermutation in the V segment. The resulting sequences were sufficient to allow annotation of the V, D and J genes, and the N1 and N2 regions constituting each unique *CDR3* and the translation of the encoded CDR3 amino acid sequence. Gene definitions were based on annotation in accordance with the IMGT database (www.imgt.org). The set of observed biological BCR IgH CDR3 sequences was normalized to correct for residual multiplex PCR amplification bias and quantified against a set of synthetic BCR IgH CDR3 sequence analogues[53]. Data were analysed using the immunoSEQ Analyzer toolset.

Multiplex cytokine analysis

Plasma collected by Ficoll isolation from heparinized whole blood was clarified by centrifugation. Magnetic Luminex assays with the MILLIPLEX MAP Human Cytokine/Chemokine Magnetic Bead Panel (MilliporeSigma, HCYTMAG-60 K-PX30) were run according to the manufacturer's protocol. The samples were quantified on the MAGPIX xMAP Instrument (Luminex). For each sample, >50 beads were collected per analyte. The median fluorescence intensity of these beads was recorded and used for analysis with the Milliplex Analyst software using a 5P regression algorithm. In the absence of signal, the assay's lower limit of detection for each cytokine was used. Patients with DS aged 7 months to 37 years ($n = 21$) and control individuals aged 6 years to 39 years ($n = 10$) were included.

Auto-antibody analysis

Seromic profiling of autoantibodies was conducted as previously described[55]. The CDI HuProt peptide array was run according to the manufacturer's instructions, at 1/500 to avoid low-tittered cross-reactivity and using a robust blocking buffer to prevent unspecific binding. After applying patient sera, chips were washed and secondary fluorescent antibodies to quantify IgG (Cy3; 532 nm) and IgA (Cy5; 635 nm) antibody isotypes were applied at the manufacturer's recommended concentrations. Each spot on the array, representing a protein printed in duplicate, was gated

using Genepix software alignment and then manually quality controlled to ensure proper quantification and removal of possible artifacts. High differences between replicate spots (CV > 0.5) were flagged, along with information about staining artifacts (rare and excluded). Raw sample-peptide matrices were then read into the R (v.4.0.4) statistical environment to be processed by the limma microarray analysis suite (v.3.46.0)[56]. On a per-sample and a per-isotype (IgG or IgA) basis, the background intensity was subtracted using the backgroundCorrect function (method: normexp, normexp.method: rma) and then normalized across all samples using the normalizeBetweenArrays function (method="cyclicloess"). Next, all control and non-detected peptides were filtered from the sample–peptide matrices and duplicate spot intensities were averaged across all of the remaining peptides on a per-sample basis. Individually processed sample matrices were concatenated, column-wise, to generate an experimental matrix containing all samples and peptides shared across all arrays and to be analysed for differential peptide reactivity using the limma differential expression functions. First, to exclude low-reactive peptides and outlier reactivity, peptides were filtered to include only peptides that had reactivities above the median signal intensity in at least 5 samples (of the 50 total samples processed). PCA was then performed using the base R (v.4.0.4) prcomp function (scale: TRUE) and PCA and loadings were visualized using the fviz_pca_ind function from the FactoMiner/factoextra package (v.1.0.7)[57]. To determine differentially abundant autoantibodies, we first fit a linear model for each peptide to model the variation of peptide reactivity (against an auto-antibody) across all arrays using limma's lmFit function. After building the model, we generated a contrast matrix to test disease groups against healthy controls. Differential autoantibody abundance testing was performed and moderated under the empirical Bayes framework described by the developer and resulting peptide lists were filtered against an absolute $\log_2$[fold change] > 1 and a $P < 0.05$. Overlap analyses were conducted and visualized using the UpSetR (v.1.4.0)[58] and ggvennDiagram (https://github.com/gaospecial/ggVennDiagram) (v.1.2.1) packages. Heat maps were generated using the ComplexHeatmap (v.2.7.4)[59] library.

Chromosomal reactivity heatmap

To assess whether autoantibody reactivity was influenced by the gene dosage effect in DS, filtered differentially abundant autoantibodies (obtained from the DS versus HC contrast; see Supplementary Data 1) were plotted across the 23 human chromosomes (hg38) in a location-dependent manner. In brief, gene names were extracted from their associated peptides and chromosomal locations were identified using the org.Hs.eg.db (v.3.12.0)[60] database. Hits were further filtered against a fold-change of 1.5 for plotting purposes and the karyoplotR (v.1.16.0)[61] package was then used for visualization.

Tissue-protein expression heat map

Tissue-protein detection values of all human proteins were extracted from the Human Protein Atlas API using the HPAanalyze (v.1.8.1)[62] package. Differentially abundant autoantibodies (obtained from the DS versus HC contrast; Supplementary Data 1) were subset from the full dataset (Supplementary Data 2) and visualized using the ComplexHeatmap library.

GO analysis

GO enrichment analysis was conducted on over-abundant ($\log_2$[fold change] > 1) autoantibodies with limma's goana function, and biological process GO terms with $P < 0.0001$ were subset and visualized using the ggplot2[63] (v.3.3.3) package.

ELISA assays

Recombinant protein ELISA

Overnight, 96-well Costar plates were coated at 4 °C with 100 µl per well of a 1 µg ml^{-1} solution of recombinant IFNGR2, MSTN or ATP6V1G2 protein (OriGene) suspended in 1× PBS. The next morning, the coating solution was removed and wells were washed three times with 100 µl of washing buffer (PBS with 0.05% (v/v) Tween 20; PBS-T). Next, 200 µl of blocking buffer (PBS with 1% with bovine plasma albumin (endotoxin-free)) was added to

each well at room temperature and incubated at room temperature for 1 h. Plasma samples were diluted 1:100 in blocking buffer.

Total IgG and 9G4 IgG ELISA

Overnight, 96-well Costar plates were coated at 4 °C with 100 µl per well of a 2 µg ml^{-1} solution of goat anti-human IgG, $F(ab')_2$-fragment-specific (Jackson ImmunoResearch) or 1:100 dilution of anti-human 9G4 IgG antibody (provided by J. Farmer) suspended in 1× PBS. The next morning, the coating solution was removed and wells were washed with three times with 100 µl of washing buffer (PBS with 0.05% (v/v) Tween-20; PBS-T). Next, 200 µl of blocking buffer (PBS with 1% with bovine plasma albumin (endotoxin-free)) was added to each well at room temperature and incubated at room temperature for 1 h. The supernatants from B cell stimulation experiments were diluted 1:10.

Antibody detection

Plates were then washed three times before adding goat anti-human IgG F(ab)–horseradish peroxidase (HRP)-conjugated secondary antibody (diluted 1:3,000 in PBS-T). The plates were incubated for 1 h at 37 °C and washed three times. The plates were again washed three times with wash buffer. Once completely dry, 100 µl SIGMAFAST OPD (*o*-phenylenediamine dihydrochloride; Sigma–Aldrich) solution was added to each well. This substrate was left on the plates for 5 min and the reaction was then stopped by the addition of 50 µl per well of 3 M hydrochloric acid. The optical density at 490 nm was measured using the POLARstar Omega (BMG Labtech) plate reader.

Detection of anti-type I IFN autoantibodies

Detection of anti-type I IFN autoantibodies was performed using Gyros as described previously[64]. Cytokines, recombinant human IFNα2 (Miltenyi Biotec, 130-108-984) or recombinant human IFNω (Merck, SRP3061) were first biotinylated with EZ-Link Sulfo-NHS-LC-Biotin (Thermo Fisher Scientific, A39257), according to the manufacturer’s instructions, with a

biotin-to-protein molar ratio of 1:12. The detection reagent contained a secondary antibody (Alexa Fluor 647 goat anti-human IgG (Thermo Fisher Scientific, A21445)) diluted in Rexxip F (Gyros Protein Technologies, P0004825; 1:500 dilution of the 2 mg ml^{-1} stock to yield a final concentration of 4 μg ml^{-1}). Phosphate-buffered saline, 0.01% Tween-20 (PBS-T) and Gyros Wash buffer (Gyros Protein Technologies, P0020087) were prepared according to the manufacturer's instructions. Plasma or serum samples were then diluted 1:100 in 0.01% PBS-T and tested with the Bioaffy 1000 CD (Gyros Protein Technologies, P0004253) and the Gyrolab xPand (Gyros Protein Technologies, P0020520). Cleaning cycles were performed in 20% ethanol.

Functional evaluation of anti-type I IFNs autoantibodies using luciferase reporter assays

The blocking activity of anti-IFNα2 and anti-IFNω autoantibodies was determined with a reporter luciferase activity as described previously[64]. HEK293T cells were transfected with a plasmid containing the firefly luciferase gene under the control of the human ISRE promoter in the pGL4.45 backbone and a plasmid constitutively expressing *Renilla* luciferase for normalization (pRL-SV40). Cells were transfected in the presence of the X-tremeGENE9 transfection reagent (Sigma-Aldrich, 6365779001) for 24 h. Cells in Dulbecco's modified Eagle medium (DMEM; Thermo Fisher Scientific) supplemented with 2% fetal calf serum and 10% healthy control or patient serum/plasma (after inactivation at 56 °C for 20 min) were stimulated with IFNα2 (Miltenyi Biotec, 130-108-984) and IFNω (Merck, SRP3061) at 10 ng ml^{-1} or 100 pg ml^{-1} for 16 h at 37 °C. Each sample was tested once for each cytokine and dose. Finally, cells were lysed for 20 min at room temperature, and luciferase levels were measured with the Dual-Luciferase Reporter 1000 Assay System (Promega, E1980) according to the manufacturer's protocol. Luminescence intensity was measured with a VICTOR-X Multilabel Plate Reader (PerkinElmer Life Sciences). Firefly luciferase activity values were normalized against *Renilla* luciferase activity values.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data supporting the findings of this study are available in the Article and its Supplementary Information.

References

1. Bull, M. J. Down syndrome. *N. Engl. J. Med.* **382**, 2344–2352 (2020).

2. Ryan, C., Vellody, K., Belazarian, L. & Rork, J. F. Dermatologic conditions in Down syndrome. *Pediatr. Dermatol.* **38**, 14731 (2021).

3. Down, J. L. H. Observations on an ethnic classification of idiots. *Lond. Hosp. Rep.* **3**, 259–262 (1866).

4. Mai, C. T. et al. National population-based estimates for major birth defects, 2010–2014. *Birth Defects Res.* **111**, 1420–1435 (2019).

5. Antonarakis, S. E. et al. Down syndrome. *Nat. Rev. Dis. Prim.* **6**, 9 (2020).

6. Rivelli, A. et al. Prevalence of endocrine disorders among 6078 individuals with Down syndrome in the United States. *J. Patient Cent. Res. Rev.* **9**, 70–74 (2022).

7. Kong, X.-F. et al. Three copies of four interferon receptor genes underlie a mild type I interferonopathy in Down syndrome. *J. Clin. Immunol.* **40**, 807–819 (2020).

8. Sullivan, K. D. et al. Trisomy 21 consistently activates the interferon response. *eLife* **5**, e16220 (2016).

9. Malle, L. et al. Excessive negative regulation of type I interferon disrupts viral control in individuals with Down syndrome. *Immunity* **55**,

2074–2084 (2022).

10. Giménez-Barcons, M. et al. Autoimmune predisposition in Down syndrome may result from a partial central tolerance failure due to insufficient intrathymic expression of AIRE and peripheral antigens. *J. Immunol.* **193**, 3872–3879 (2014).

11. Araya, P. et al. Trisomy 21 dysregulates T cell lineages toward an autoimmunity-prone state associated with interferon hyperactivity. *Proc. Natl Acad. Sci. USA* **116**, 24231–24241 (2019).

12. Cossarizza, A. et al. Age-related expansion of functionally inefficient cells with markers of natural killer activity in Down's syndrome. *Blood* **77**, 1263–1270 (1991).

13. Verstegen, R. H. J. & Kusters, M. A. A. Inborn errors of adaptive immunity in Down syndrome. *J. Clin. Immunol.* **40**, 791–806 (2020).

14. Taft, J. & Bogunovic, D. The Goldilocks zone of type I IFNs: lessons from human genetics. *J. Immunol.* **201**, 3479–3485 (2018).

15. Rea, I. M. et al. Age and age-related diseases: role of inflammation triggers and cytokines. *Front. Immunol.* **9**, 586 (2018).

16. Zhang, Y. et al. Aberrations in circulating inflammatory cytokine levels in patients with Down syndrome: a meta-analysis. *Oncotarget* **8**, 84489–84496 (2017).

17. Huggard, D. et al. Increased systemic inflammation in children with Down syndrome. *Cytokine* **127**, 154938 (2020).

18. Junttila, I. S. Tuning the cytokine responses: an update on interleukin (IL)-4 and IL-13 receptor complexes. *Front. Immunol.* **9**, 888 (2018).

19. Lambert, K. et al. Deep immune phenotyping reveals similarities between aging, Down syndrome, and autoimmunity. *Sci. Transl. Med.* **14**, 4888 (2022).

20. Milner, J. D. et al. Early-onset lymphoproliferation and autoimmunity caused by germline STAT3 gain-of-function mutations key points. *Blood* **21**, 25 (2015).

21. Fabre, A. et al. Clinical aspects of STAT3 gain-of-function germline mutations: a systematic review. *J. Allergy Clin. Immunol. Pract.* **7**, 1958–1969 (2019).

22. Verstegen, R. H. J. et al. Defective B-cell memory in patients with Down syndrome. *J. Allergy Clin. Immunol.* **134**, 1346–1353 (2014).

23. Dieudonné, Y. et al. Immune Defect in adults with down syndrome: insights into a complex issue. *Front. Immunol.* **11**, 840 (2020).

24. Song, W. et al. Development of Tbet- and CD11c-expressing B cells in a viral infection requires T follicular helper cells outside of germinal centers. *Immunity* **55**, 290–307 (2022).

25. Karnell, J. L. et al. Role of $CD11c^{+}$ $T\text{-}bet^{+}$ B cells in human health and disease. *Cell. Immunol.* **321**, 40–45 (2017).

26. Woodruff, M. C. et al. Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19. *Nat. Immunol.* **21**, 1506–1516 (2020).

27. Jenks, S. A. et al. Distinct effector B cells induced by unregulated toll-like receptor 7 contribute to pathogenic responses in systemic lupus erythematosus. *Immunity* **49**, 725–739 (2018).

28. Keller, B. et al. The expansion of human $T\text{-}bet^{high}CD21^{low}$ B cells is T cell dependent. *Sci. Immunol.* **6**, 52 (2021).

29. Li, H., Borrego, F., Nagata, S. & Tolnay, M. Fc receptor-like 5 expression distinguishes two distinct subsets of human circulating tissue-like memory B cells. *J. Immunol.* **196**, 4064–4074 (2016).

30. Sokol, C. L. & Luster, A. D. The chemokine system in innate immunity. *Cold Spring Harb. Perspect. Biol.* **7**, a016303 (2015).

31. Vestergaard, C. et al. A Th2 chemokine, TARC, produced by keratinocytes may recruit CLA^+CCR4^+ lymphocytes into lesional atopic dermatitis skin. *J. Invest. Dermatol.* **115**, 640–646 (2000).

32. Mikhak, Z., Strassner, J. P. & Luster, A. D. Lung dendritic cells imprint T cell lung homing and promote lung immunity through the chemokine receptor CCR4. *J. Exp. Med.* **210**, 1855 (2013).

33. Zumaquero, E. et al. IFNγ induces epigenetic programming of human T-bet^{hi} B cells and promotes TLR7/8 and IL-21 induced differentiation. *eLife* **8**, e41641 (2019).

34. Wang, S. et al. IL-21 drives expansion and plasma cell differentiation of autoreactive CD11c^{hi} T-bet + B cells in SLE. *Nat. Commun.* **9**, 1758 (2018).

35. Jego, G. et al. Plasmacytoid dendritic cells induce plasma cell differentiation through type I interferon and interleukin 6. *Immunity* **19**, 225–234 (2003).

36. Naradikian, M. S. et al. Cutting edge: IL-4, IL-21, and IFN-γ interact to govern T-bet and CD11c expression in TLR-activated B cells. *J. Immunol.* **197**, 1023–1028 (2016).

37. Stone, S. L. et al. T-bet transcription factor promotes antibody-secreting cell differentiation by limiting the inflammatory effects of IFN-γ on B cells. *Immunity* **50**, 1172–1187 (2019).

38. Bashford-Rogers, R. J. M. et al. Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature* **574**, 122–126 (2019).

39. Tipton, C. M., Hom, J. R., Fucile, C. F., Rosenberg, A. F. & Sanz, I. Understanding B-cell activation and autoantibody repertoire selection in systemic lupus erythematosus: a B-cell immunomics approach. *Immunol. Rev.* **284**, 120–131 (2018).

40. Richardson, C. et al. Molecular basis of 9G4 B cell autoreactivity in human systemic lupus erythematosus. *J. Immunol.* **191**, 4926–4939 (2013).

41. Marcovecchio, G. E. et al. Thymic epithelium abnormalities in DiGeorge and Down syndrome patients contribute to dysregulation in T cell development. *Front. Immunol.* **10**, 447 (2019).

42. Pierce, M. J., LaFranchi, S. H. & Pinter, J. D. Characterization of thyroid abnormalities in a large cohort of children with Down syndrome. *Horm. Res. Paediatr.* **87**, 170 (2017).

43. Mader, S., Brimberg, L. & Diamond, B. The role of brain-reactive autoantibodies in brain pathology and cognitive impairment. *Front. Immunol.* **8**, 1101 (2017).

44. Lepagnol-Bestel, A. M. et al. DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. *Hum. Mol. Genet.* **18**, 1405–1414 (2009).

45. van der Poel, C. E., Spaapen, R. M., van de Winkel, J. G. J. & Leusen, J. H. W. Functional characteristics of the high affinity IgG receptor, FcγRI. *J. Immunol.* **186**, 2699–2704 (2011).

46. Bustamante, J., Boisson-Dupuis, S., Abel, L. & Casanova, J. L. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-γ immunity. *Semin. Immunol.* **26**, 454–470 (2014).

47. Malle, L. et al. Individuals with Down syndrome hospitalized with COVID-19 have more severe disease. *Genet. Med.* **23**, 576–580 (2020).

48. Gillespie, K. M. et al. Islet autoimmunity in children with Down's syndrome. *Diabetes* **55**, 3185–3188 (2006).

49. Kohen, D. & Wise, P. H. Autoantibodies in Down's syndrome. *Lancet* **340**, 430 (1992).

50. Cardinale, K. M. et al. Immunotherapy in selected patients with Down syndrome disintegrative disorder. *Dev. Med. Child Neurol.* **61**, 847–851 (2019).

51. Geanon D, et al. A streamlined whole blood CyTOF workflow defines a circulating immune cell signature of COVID-19. *Cytometry A.* **99**, 446-461 (2021).

52. Robins, H. S. et al. Comprehensive assessment of T-cell receptor β-chain diversity in αβ T cells. *Blood* **114**, 4099–4107 (2009).

53. Carlson, C. S. et al. Using synthetic templates to design an unbiased multiplex PCR assay. *Nat. Commun.* **4**, 2680 (2013).

54. Robins, H. et al. Ultra-sensitive detection of rare T cell clones. *J. Immunol. Methods* **375**, 14–19 (2012).

55. Gnjatic, S. et al. Seromic profiling of ovarian and pancreatic cancer. *Proc. Natl Acad. Sci. USA* **107**, 5088–5093 (2010).

56. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

57. Lê, S., Josse, J. & Husson, F. FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).

58. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).

59. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

60. Carlson, M. org.Hs.eg.db (2017).

61. Gel, B. karyoploteR (2017).

62. Nhat, A. HPAanalyze (2018).

63. Wickham, H. ggplot2 (2009); https://doi.org/10.1007/978-0-387-98141-3.

64. Bastard, P. et al. Autoantibodies neutralizing type I IFNs are present in ~4% of uninfected individuals over 70 years old and account for ~20% of COVID-19 deaths. *Sci. Immunol.* **6**, abl4340 (2021).

Acknowledgements

We thank all of the patients and their families for their participation; A. Rahman, D. Geanon and G. Kelly from the Human Immune Monitoring Centre at the Icahn School of Medicine for their technical assistance; and J. Farmer and K. Hillier for sharing reagents. This study was funded by the National Institute of Allergy and Infectious Diseases Grants R01AI150300, R01AI150300-01S1 and R01AI151029. L. Malle was supported by the National Institute of Child Health and Human Development T32 training grant T32HD075735. L.N. is supported by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, NIH. S.G. was supported by NIH grants CA224319, DK124165 and CA196521.

Author information

Author notes

1. These authors contributed equally: Marta Martin-Fernandez, O'Jay Stewart

Authors and Affiliations

1. Center for Inborn Errors of Immunity, Icahn School of Medicine at Mount Sinai, New York, NY, USA

   Louise Malle, Roosheel S. Patel, Marta Martin-Fernandez, O Jay Stewart, Sofija Buta, Ashley Richardson, Justin Taft & Dusan Bogunovic

2. Department of Pediatrics, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Louise Malle, Roosheel S. Patel, Marta Martin-Fernandez, O Jay Stewart, Sofija Buta, Ashley Richardson, Justin Taft, Rebecca Trachtman, Douglas Bush & Dusan Bogunovic

3. Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Louise Malle, Roosheel S. Patel, Marta Martin-Fernandez, O Jay Stewart, Sofija Buta, Ashley Richardson, Justin Taft, Julie Samuels, Rebecca Trachtman & Dusan Bogunovic

4. Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Louise Malle, Roosheel S. Patel, Marta Martin-Fernandez, O Jay Stewart, Sofija Buta, Ashley Richardson, Vanessa Barcessat, Justin Taft, Kevin Tuballes, Sacha Gnjatic & Dusan Bogunovic

5. Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Louise Malle, Roosheel S. Patel, Marta Martin-Fernandez, O Jay Stewart, Sofija Buta, Ashley Richardson, Justin Taft, Brad R. Rosenberg & Dusan Bogunovic

6. Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U1163, Necker Hospital for Sick Children, Paris, France

Quentin Philippot, Paul Bastard & Jean-Laurent Casanova

7. University of Paris, Imagine Institute, Paris, France

Quentin Philippot, Paul Bastard & Jean-Laurent Casanova

8. St Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, NY, USA

Paul Bastard & Jean-Laurent Casanova

9. Pediatric Hematology-Immunology and Rheumatology Unit, Necker Hospital for Sick Children, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

   Paul Bastard

10. Institut Jérôme Lejeune, Paris, France

    Clotilde Mircher, Anne-Sophie Rebillat, Louise Maillebouis & Marie Vilaire-Meunier

11. Department of Pediatrics, Necker Hospital for Sick Children, Paris, France

    Jean-Laurent Casanova

12. Howard Hughes Medical Institute, New York, NY, USA

    Jean-Laurent Casanova

13. Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

    Luigi D. Notarangelo

Contributions

L. Malle designed and performed experiments and analysed the data for cytokine array, CyTOF, B–T cell co-cultures and BCR sequencing, and wrote the manuscript. R.S.P. analysed the CDI array and edited the manuscript. M.M.-F. performed steady-state B–T cell co-cultures from DS donors. O.S. performed 9G4 and IFNy ELISA. J.T. edited the manuscript. Q.P. performed IFN autoantibody Gyros and neutralization assays. S.B. and A.R. coordinated DS cohort and processed whole-blood samples. V.B., K.T. and S.G. performed the CDI array. B.R.R. helped to analyse BCR sequencing data. P.B., J.S., C.M., A.-S.R., L. Maillebouis,

M.V.-M., R.T., J.-L.C., L.N. and D. Bush recruited patients. D. Bogunovic supervised the work, wrote the manuscript, and helped to design the experiments and analyse the data.

Corresponding author

Correspondence to Dusan Bogunovic.

Ethics declarations

Competing interests

D.B. is the founder and part owner of Lab11 Therapeutics. S.G. reports other research funding from Genentech, Boehringer-Ingelheim, Celgene, Takeda, and Regeneron.

Peer review

Peer review information

Nature thanks Stuart Tangye and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Global cytokine dysregulation in DS.

(**A**-**B**) Correlation between cytokine group and (**A**) Age and **(B**) Clinical Immune Dysfunction score (determined according to reported clinical history, see Methods) and Age in individuals with DS. Significance assessed

by one-way ANOVA with Tukey's post-hoc analysis, ns denotes $p > 0.05$, $*p \leq 0.05$; $**p \leq 0.005$, $***p \leq 0.0005$. (**C**) Multiplex cytokine analysis by Magnetic Luminex assay of plasma from HCs (n = 10), uninfected individuals with DS (n = 21), and individual with acute respiratory infection (n = 1), or individuals with DS and COVID-19 (n = 7) with severity and time of sampling as indicated, expressed as log2FC over the mean HC per cytokine. Unsupervised clustering of samples and cytokines using the complete method (distance metric: Euclidean). (**D**) Intracellular staining of IL-6 in CD4 T cells, CD8 T cells, and Myeloid cells from HCs (n = 4) and individuals with DS (n = 5).

Extended Data Fig. 2 Cellular immune landscape in DS.

(**A**) Representative t-SNE of agranulocytes adults with DS (n = 3) and age-matched HCs (n = 3) illustrating the immune cell distribution in whole blood. (**B**-**E**) Frequencies of (**B**) agranulocyte subsets, (**C**) granulocyte subsets, and (**D**) CD4 and (**E**) CD8 T cell subsets in whole blood from adults with DS (n = 10), adults with DS and COVID-19 (n = 5), and age-matched HCs (n = 8). **(F-G**) Frequencies (**F**) and absolute counts (**G**) of T regulatory (Treg) cells ($CD4^{+}CD25^{+}CD127^{-}$) and naive ($CD45RA^{+}$) and memory ($CD45RA^{-}$) subsets in whole blood from adults with DS (n = 5) and age-matched HCs (n = 5). (**H-I**) Representative plots and calculated frequencies of (**H**) T helper (Th) and (**I**) T follicular helper (Tfh) cell subsets in whole blood from adults with DS (*n* = *5*) and age-matched HCs (*n* = 5), expressed as percent of memory CD4 T cells. In (B-E), whiskers denote min and max values, bounds of box denote Q1–Q3, and centre bar denotes mean. (F-I) Error bars denote SD. (B-E) Significance assessed by one-way ANOVA with Tukey's post-hoc analysis, ns denotes $p > 0.05$, $*p \leq 0.05$; $**p \leq 0.005$, $***p \leq 0.0005$. (F-I) Significance assessed by two-tailed paired t tests, ns denotes $p > 0.05$, $*p \leq 0.05$; $**p \leq 0.005$, $***p \leq 0.0005$.

Extended Data Fig. 3 Basal signalling in CD4 T cells in DS.

(**A**) Basal STAT3 phosphorylation in CD4 T cell subsets from HCs (n = 13), individuals with DS (n = 19), and adults with DS and COVID-19 (n = 5) expressed as Log2FC over the mean HCs per subset. Significance assessed by one-way ANOVA with Tukey's post-hoc analysis, ns denotes $p > 0.05$,

*p ≤ 0.05; **p ≤ 0.005, ***p ≤ 0.0005. Whiskers denote min and max values, bounds of box denote Q1–Q3, and centre bar denotes mean. (**B**) Basal STAT5 phosphorylation in CD4 T cell subsets from individuals with DS (n = 14) and age-matched controls (n = 10), expressed as Log2FC over the mean HCs per subset. (**C**-**D**) STAT3 phosphorylation in CD4 T cell subsets expressed as Log2FC over the mean mock-treated HCs per subset after ex vivo whole blood treatment for 4 h with (**C**) Tofacitinib (500nM) (n = 6 HC, n = 7 DS), (**D**) Tocilizumab (50 µg ml^{-1}) (n = 4 HC, n = 7 DS). (**E**) Surface expression of IL-6R in CD4 T cells from adults with DS (n = 3) and age-matched HCs (n = 3). (**F**) STAT3 phosphorylation in CD4 T cell subsets induced by stimulation of whole blood with recombinant IL-6 (50 ng ml^{-1}) for 15 min, expressed as Log2FC over the mean mock-treated HCs (n = 2 HC, n = 4 DS). Box plots denote min and max values for HCs, error bars indicate SD and centre denotes mean for DS. (C-E) Error bars denote SD. (B-F) Significance assessed by two-tailed paired t tests, ns denotes p > 0.05, *p ≤ 0.05; **p ≤ 0.005, ***p ≤ 0.0005.

Extended Data Fig. 4 Abnormal B cell subsets in DS.

(**A**) Raw B cell counts in HCs (n = 6) and adults with DS (n = 7) or patients with SLE (n = 4). Error bars denote SD. (**B**) Gating scheme of B cells subtypes. (**C**) Frequency in children with DS (n = 11) and age-matched HCs (n = 7) of total B cells expressed as percent of $CD66b^{-}$ cells (non-granulocytes), Significance assessed by two-tailed unpaired t tests, *p ≤ 0.05. (**D**-**F**) Frequency in children with DS (n = 8) and age-matched HCs (n = 4) of $CD11c^{+}$ B cells in the IgD^{+} naive or DN compartments, expressed as percent of (**D**) total, (**E**) IgD^{+} naïve, or (**F**) DN B cells. (**G**) Ratios of $CD11c^{+}$ subsets to $CD11c^{-}$ subset (rN:aN and DN2:DN1) in HC and DS groups, Log2-transformed. Significance assessed by two-tailed unpaired t tests, *p ≤ 0.05. (**H**-**I**) Correlation of $CD11c^{+}$ B cell frequency and (**H**) age and (**I**) total cytokines (calculated as the sum of all circulating cytokines, pg ml^{-1}) in individuals (H) or adults (I) with DS and age-matched controls. (r: Pearson correlation coefficient). (**J**) Correlation of $CD11c^{+}$ B cell frequency and autoimmune score (calculated as the sum of autoimmune diseases) in individuals with DS. (r: Pearson correlation coefficient). (**K**) Total IgG in the plasma of HCs (n = 7) and individuals with DS (n = 12) or SLE (n = 6) assessed by ELISA. Significance assessed by one-way ANOVA with

Tukey's post-hoc analysis, ns denotes $p > 0.05$, $*p \leq 0.05$. In (C-F, K), whiskers denote min and max values, bounds of box denote Q1–Q3, and centre bar denotes mean.

Extended Data Fig. 5 Atypical activation of naive B cells by cytokines and activated T cells.

(**A** Plasmablast differentiation of MACS-isolated total B cells from 2 healthy donors after 3-day culture in the presence of BAFF (10 ng ml^{-1}), IL-2 (50 ng ml^{-1}), IL-21 (50 ng ml^{-1}), R848 (1 µg ml^{-1}), anti-IgM (1 µg ml^{-1}), and IgG-depleted plasma from HCs ($n = 3$) or individuals with DS ($n = 3$), run in duplicates. (**B**-**C**) MACS-isolated total B cells from 2 healthy donors were cultured in the presence of BAFF, IL-2, IL-21, R848, anti-IgM (same concentrations as A), in the presence of IL-6 (100 ng ml^{-1}), IFN-α2b (100 U ml^{-1}), or both, run in duplicates. After 3 days, cells were washed and cultured for another 3 days in media with anti-IgM. (**B**) Plasmablast differentiation at day 3 and (**C**) ELISA for total IgG in supernatant at day 6. (**D**-**E**) MACS-isolated total B cells from 2 healthy donors were cultured for 3-days in the presence of BAFF, IL-2, IL-21, R848, anti-IgM (same concentrations as A), and IL-4 (20 ng ml^{-1}), IFN-ɣ (20 ng ml^{-1}), or both, run in duplicates. Cells were then washed and cultured for another 3 days in media with anti-IgM. (**D**) Plasmablast differentiation at day 3 and (**E**) ELISA for total IgG in supernatant at day 6. (**F**) Plasmablast differentiation of MACS-isolated naive B cells from a healthy donor after 3-day culture in the presence of BAFF, IL-2, IL-21, R848, anti-IgM (same concentrations as J), and IgG-depleted DS plasma in the presence of Tofacitinib (500nM) or antibodies blocking IFN-I, IFN-II, IL-6, and TFN-ɑ signalling, run in duplicates. (**G**-**H**) Co-cultures containing T cells activated with IL-6, IL-2, both, or polarized into "Th1 cells" together with MACS-isolated naive B cells from the same donor, run in triplicates. (**G**) Intracellular T-bet expression in non-plasmablast B cells and (**H**) quantification of IFN-g in the supernatant after 3-6 days of co-culture. (**I**) Frequency of plasmablasts in co-cultures containing T cells previously polarized with serum from HCs ($n = 3$) or individuals with DS ($n = 3$) together with MACS-isolated naive B cells from the same donor. Significance assessed by two-tailed unpaired t-test. (A-I) Error bars denote SD. (A,H) Significance assessed by One-way ANOVA

with Tukey's post-hoc analysis, ns denotes $p > 0.05$; **$p \leq 0.005$; ***$p \leq 0.0005$; ****$p \leq 0.0001$.

Extended Data Fig. 6 Receptor sequencing in atypical B cells.

(**A**) Expression of IgD and IgA in CD11c$^+$ B cells from adults with DS ($n = 3$) and age-matched HCs ($n = 3$). Significance assessed by unpaired t-tests, ns denotes $p > 0.05$. Error bars denote SD. (**B**-**F**) BCR sequencing from gDNA isolated from sorted naïve, CD11c$^+$ and memory B cells from controls ($n = 6$) and individuals with DS ($n = 6$). (**B**) Sorting scheme for naïve, CD11c$^+$ and memory B cells (left) and number of cells sorted and number of productive BCRs obtained by sequencing for each cell type (right). (**C**) Fraction of in-frame BCRs containing no stop codons ("productive BCRs") in naïve, CD11c$^+$ and memory B cells from controls ($n = 6$) and individuals with DS ($n = 6$). (**D**) Simpson clonality in productive BCRs from CD11c$^+$ B cells from controls ($n = 6$) and individuals with DS ($n = 6$). In (C-D), significance assessed by two-tailed unpaired t-test (ns denotes $p > 0.05$) and whiskers denote min and max values, bounds of box denote Q1–Q3, and centre bar denotes mean. (**E**-**F**) Representative heatmaps of Morisita index indicating overlap between B cell subsets in HC and DS at the (**E**) nucleotide and (**F**) amino acid levels. (**G**) Heatmap of IGHV gene frequency in HC and DS. Only genes that occurred at higher than 0.1% frequency in more than 10 samples are shown. (**H**) 9G4 IgG antibodies in supernatant after 4-day culture of sorted HC naïve, CD11c$^+$ or memory B cells in the presence of BAFF, IL-2, IL-10, IL-21, the TLR7/8 ligand R848 with or without IFN-ɣ. Results representative of 2 independent experiments.

Extended Data Fig. 7 Autoantibodies in DS plasma.

(**A**) Principal component analysis of HuProt IgA dataset for adult HCs ($n = 4$), adults with DS ($n = 5$), and patients, Immunodysregulation polyendocrinopathy enteropathy X-linked syndrome (IPEX) ($n = 3$) and Autoimmune polyglandular syndrome type 1 (APS-1) ($n = 1$). (**B**) Number of IgA autoantigens enriched at least 2-fold in DS, IPEX, and APS-1 compared to HCs. (**C**) Venn diagram of enriched IgA autoantigens overlapping between disease groups. (**D**) Heatmap of enriched IgA autoantigens in HCs, DS, IPEX, and APS-1. Colour intensity corresponds to

the log2FC expression value relative to the mean of healthy adult controls. (**E**) Component loadings (PC1 and PC2) from Principal component analysis of HuProt IgG dataset for adults with DS ($n = 5$) and age-matched HCs ($n = 4$). (**F**) ELISA of anti-MSTN and anti-ATP6V1G2 autoantibodies in plasma from HCs ($n = 6$) and individuals with DS ($n = 6$). Significance assessed by two-tailed unpaired t tests, * denotes $p \leq 0.05$. Whiskers denote min and max values, bounds of box denote Q1–Q3, and centre bar denotes mean. (**G**) Heatmap of IgG type I IFN autoantigens in HCs, DS, IPEX, and APS-1. Colour intensity corresponds to the log2FC expression value relative to the mean of healthy adult controls. (**H**) Quantification of antibodies against IFNα2 and IFNω in serum from HCs ($n = 32$), DS without COVID-19 ($n = 10$), DS with COVID-19 ($n = 6$), and APS-1 ($n = 1$) by Gyros assay. (**I**) Neutralization of IFNα2 and IFNω activity by serum from HCs ($n = 32$), DS without COVID-19 ($n = 10$), DS with COVID-19 ($n = 6$), and APS-1 ($n = 1$) by Gyros assay. (H-I) Significance assessed by One-way ANOVA with Tukey's post-hoc analysis, ns denotes $p > 0.05$, $*p \leq 0.05$; $**p \leq 0.005$; $***p \leq 0.0005$; $****p \leq 0.0001$.

Extended Data Fig. 8 Graphical abstract.

Breaking Immune Tolerance in Down Syndrome: A Triad of Cytokines, Activated T cells and CD11c+ B Cells. Created with BioRender.

Extended Data Table 1 Clinical history and calculated immune dysfunction scores of individuals with DS
Extended Data Table 2 Clinical summary of the individuals with DS studied
Extended Data Table 3 Immunological and clinical features of individuals with SLE studied

Supplementary information

Supplementary Figure 1

Schematic of the gating hierarchy used to define immune populations by mass cytometry.

Reporting Summary

Supplementary Data 1

Differential peptide analysis of HuProt IgG dataset.

Supplementary Data 2

HuProt raw data files.

Rights and permissions

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Reprints and Permissions

About this article

Cite this article

Malle, L., Patel, R.S., Martin-Fernandez, M. *et al.* Autoimmunity in Down's syndrome via cytokines, CD4 T cells and CD11c^{+} B cells. *Nature* **615**, 305–314 (2023). https://doi.org/10.1038/s41586-023-05736-y

- Received: 16 December 2021
- Accepted: 17 January 2023
- Published: 22 February 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05736-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

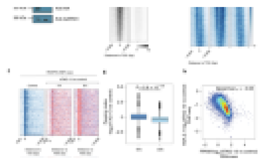
Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

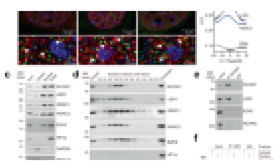
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05736-y



A dual role for H2A.Z.1 in modulating the dynamics of RNA polymerase II initiation and elongation

10 May 2021

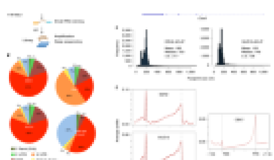
Constantine Mylonas, Choongman Lee, … Laurie A. Boyer



Unveiling RCOR1 as a rheostat at transcriptionally permissive chromatin

23 March 2022

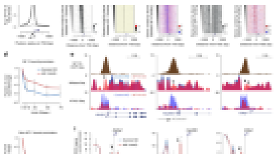
Carlos Rivera, Hun-Goo Lee, … María Estela Andrés



Motif-driven interactions between RNA and PRC2 are rheostats that regulate transcription elongation

04 January 2021

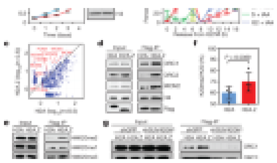
Michael Rosenberg, Roy Blum, … Jeannie T. Lee



[NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region]

11 July 2019

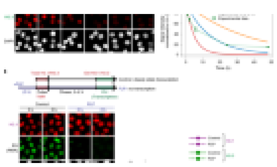
Andrew J. Oldfield, Telmo Henriques, … Raja Jothi



[H2A.Z facilitates licensing and activation of early replication origins]

25 December 2019

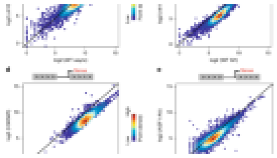
Haizhen Long, Liwei Zhang, … Guohong Li



[Two HIRA-dependent pathways mediate H3.3 de novo deposition and recycling during transcription]

07 September 2020

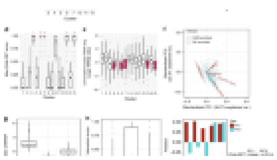
Júlia Torné, Dominique Ray-Gallet, … Geneviève Almouzni



[Distinct transcriptional roles for Histone H3-K56 acetylation during the cell cycle in Yeast]

26 September 2019

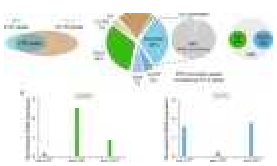
Salih Topal, Pauline Vasseur, … Craig L. Peterson



[Long ncRNA A-ROD activates its target gene DKK1 at its release from chromatin]

24 April 2018

Evgenia Ntini, Annita Louloupi, … Ulf Andersson Vang Ørom



[XPC is an RNA polymerase II cofactor recruiting ATAC to promoters by interacting with E2F1]

04 July 2018

B. Bidon, I. Iltis, … N. Le May

- Article
- Open Access
- Published: 01 March 2023

H3K4me3 regulates RNA polymerase II promoter-proximal pause-release

- Hua Wang,
- Zheng Fan,
- Pavel V. Shliaha,
- Matthew Miele,
- Ronald C. Hendrickson,
- Xuejun Jiang &
- …
- Kristian Helin

Show authors

Nature **volume 615**, pages 339–348 (2023)

- 21k Accesses
- 1479 Altmetric
- Metrics details

Abstract

Trimethylation of histone H3 lysine 4 (H3K4me3) is associated with transcriptional start sites and has been proposed to regulate transcription initiation[1,2]. However, redundant functions of the H3K4 SET1/COMPASS methyltransferase complexes complicate the elucidation of the specific role of H3K4me3 in transcriptional regulation[3,4]. Here, using mouse embryonic stem cells as a model system, we show that acute ablation of shared subunits of the SET1/COMPASS complexes leads to a complete loss of all H3K4 methylation. Turnover of H3K4me3 occurs more rapidly than that of H3K4me1 and H3K4me2 and is dependent on KDM5 demethylases.

Notably, acute loss of H3K4me3 does not have detectable effects on transcriptional initiation but leads to a widespread decrease in transcriptional output, an increase in RNA polymerase II (RNAPII) pausing and slower elongation. We show that H3K4me3 is required for the recruitment of the integrator complex subunit 11 (INTS11), which is essential for the eviction of paused RNAPII and transcriptional elongation. Thus, our study demonstrates a distinct role for H3K4me3 in transcriptional pause-release and elongation rather than transcriptional initiation.

Main

Histone modifications are closely linked to transcription regulation and are involved in processes that determine cell fate, development and disease[5,6]. Methylation of H3 lysine 4 (H3K4) is one of the most studied modifications owing to its association with gene expression and cancer[2,7]. H3K4 methylation is deposited by SET1/COMPASS complexes that contain different lysine methyltransferases and several essential subunits, including the six catalytic subunits SETD1A, SETD1B and MLL1–4. H3K4me3 is enriched at transcription start sites (TSSs) and is believed to promote transcription through the recruitment of PHD-domain-containing proteins involved in transcription initiation, such as TATA-box-binding protein associated factor 3 (TAF3)[1,8]. Moreover, H3K4me3 has been reported to counteract DNA methylation[9] and repressive histone modifications such as H3K9me3 and H3K27me3[10,11]. Furthermore, the broad H3K4me3 domains, prominently found in preimplantation embryos[10] and somatic cells[12], have been proposed to ensure transcriptional consistency of essential genes to maintain cell identity.

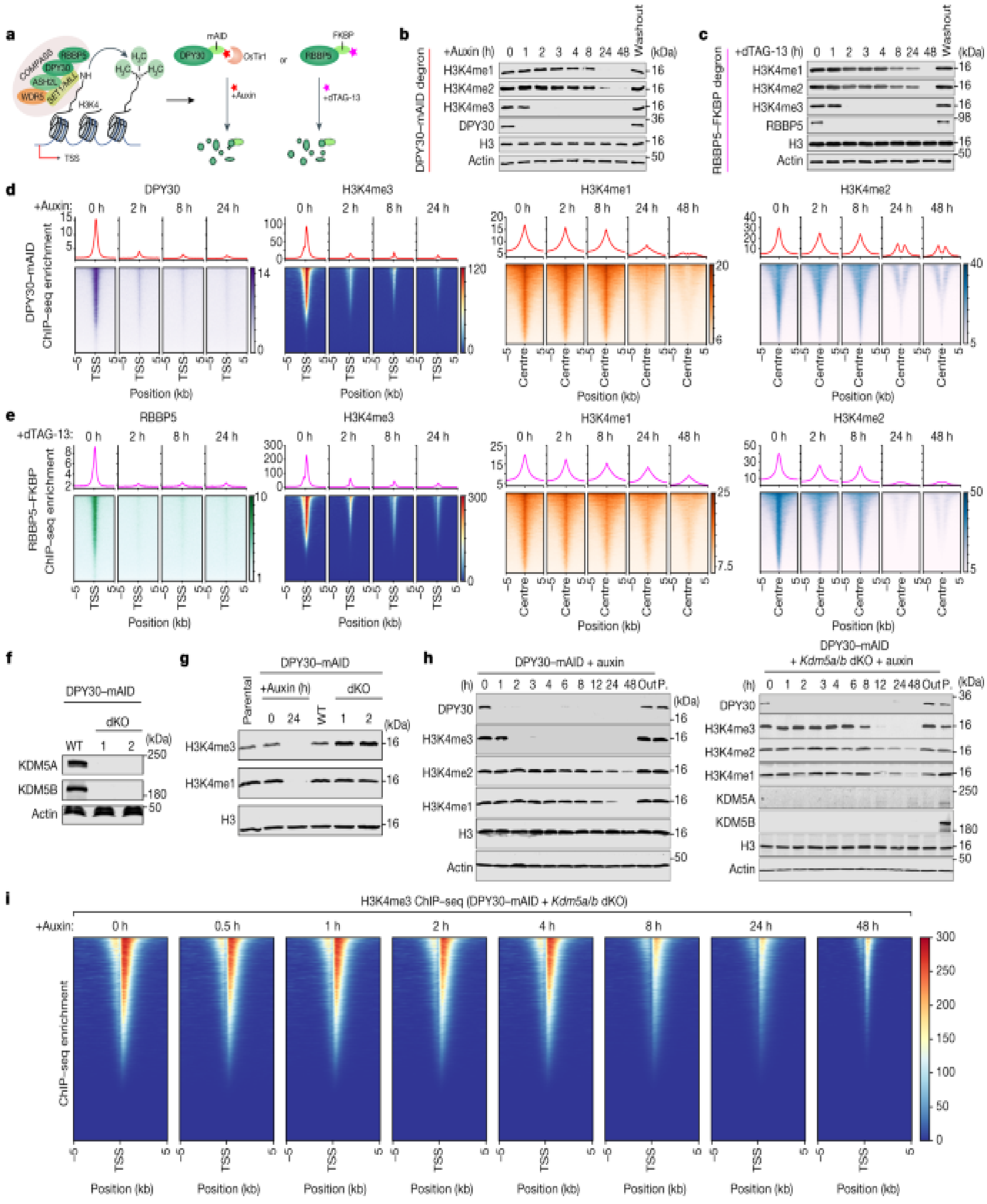
Previous studies have addressed the roles of various components of the SET1/COMPASS complexes, such as CFP1[13], WDR5[14], DPY30[15], MLL1/2[16], MLL3/4[17] and SETD1A/B[18], in mammalian cells by inhibiting their expression or deleting their respective genes. However, observations of the effects on gene expression in these studies have not produced uniform results. These inconsistencies may in part be due to redundant functions of the components of the SET1/COMPASS complexes. Moreover, inadequate temporal resolution of previous knockdown or knockout studies has

rendered it difficult to establish a direct role for H3K4 methylation in gene expression[19]. In this study, we determined the transcriptional effects of acute loss of H3K4me3 by targeted degradation of core components of the SET1/COMPASS complexes. Our data show that H3K4me3 has a key role in regulating RNAPII pause-release; however, notably, we did not detect a role for H3K4me3 in regulating transcription initiation.

Models to study H3K4 methylation

To study the role of H3K4me3 in transcription, we developed model systems for the acute depletion of either DPY30 or RBBP5, two core components that have been reported to be integral and shared components of all the SET1/COMPASS family of H3K4 methyltransferase complexes, in mouse embryonic stem (mES) cell lines (Fig. 1a and Extended Data Fig. 1a,b). To do this, we generated isogenic mES cells expressing DPY30 fused to the miniAID degron tag[20] (DPY30–mAID) or RBBP5 fused to the FKBP12(F36V) degron tag[21] (RBBP5–FKBP) (Extended Data Fig. 1c). Auxin treatment led to undetectable levels of DPY30 within 1 h, and treatment with auxin for 24 h caused a substantial decrease in mono-, di- and trimethylation of H3K4 (Fig. 1b and Extended Data Fig. 1d). Similarly, treatment of RBBP5–FKBP cells with dTAG-13 also caused acute loss of H3K4 methylation after the depletion of RBBP5 (Fig. 1c). Thus, these results show that targeting DPY30 or RBBP5 of the SET1/COMPASS complex leads to a rapid loss of H3K4 methylation.

Fig. 1: Acute depletion of SET1/COMPASS core subunits reveals rapid turnover of H3K4me3.



a, Schematic of the degron systems for the targeted degradation of DPY30 and RBBP5. **b**,**c**, Immunoblot analysis of DPY30, RBBP5 and H3K4me1–3 levels at the indicated times after treatment with 500 nM auxin (**b**) or 500 nM dTAG-13 (**c**). Washout, degron ligand was washed out for 48 h. **d**,**e**, ChIP–seq heat maps and profiles were generated from control and auxin-

treated DPY30–mAID cells (**d**) and dTAG-13-treated RBBP5–FKBP cells (**e**). For DPY30, RBBP5 and H3K4me3 ChIP–seq, the signal was plotted over the TSSs (TSS ± 5 kb) of protein-coding genes. For H3K4me1 and H3K4me2 ChIP–seq, the signal was plotted over their centre peaks (peak centre ± 5 kb), which are called from steady-state mES cells. Sites were sorted by the ChIP–seq signals at 0 h. **f**, Immunoblot analysis of KDM5A and KDM5B in DPY30–mAID cells and two independently isolated dKO cell lines. β-Actin was used as the loading control. **g**, Immunoblot analysis of H3K4me3 and H3K4me1 levels in DPY30–mAID, control and *Kdm5a/b*-dKO cells. Histone H3 was used as the loading control. **h**, Immunoblot analysis of DPY30, H3K4me1–3, KDM5A and KDM5B at the indicated times after auxin treatment. Out, degron ligand was washed out for 48 h; P, parental cells. **i**, H3K4me3 ChIP–seq heat maps in DPY30–mAID *Kdm5a/b*-dKO cells. The signal was plotted over the TSSs (TSS ± 5 kb) of protein-coding genes. Rows are sorted by decreasing ChIP–seq occupancy in the auxin 0 h cells.

Notably, H3K4me3 levels were strongly decreased within 2 h of DPY30- and RBBP5-induced degradation, whereas a substantial decrease in H3K4me1 and H3K4me2 levels was achieved only after 24 h (Fig. 1b,c). The binding of DPY30 and RBBP5 was completely lost genome-wide, as determined by chromatin immunoprecipitation with sequencing (ChIP–seq) analysis after treatment for 2 h with auxin or dTAG-13, respectively. Moreover, all three methylated forms of H3K4 were strongly decreased genome-wide after treatment for 24 h with auxin or dTAG-13 (Fig. 1d,e and Extended Data Fig. 1e–h). Consistently, H3K4me3 exhibited a global decrease on average across all TSS regions genome-wide within 2 h of auxin/dTAG-13 treatment. These effects on H3K4 methylation were further validated by analysis using ChIP with quantitative PCR (ChIP–qPCR) (Extended Data Fig. 1i–l). The loss of DPY30 and RBBP5 prevented long-term subcloning of the degron cells, whereas their acute degradation led to slower proliferation (Extended Data Fig. 2a–c).

Rapid turnover of H3K4me3 by KDM5

The rapid loss of H3K4me3 suggests that H3K4me3 levels are dynamically regulated by histone turnover and/or active demethylation. To investigate

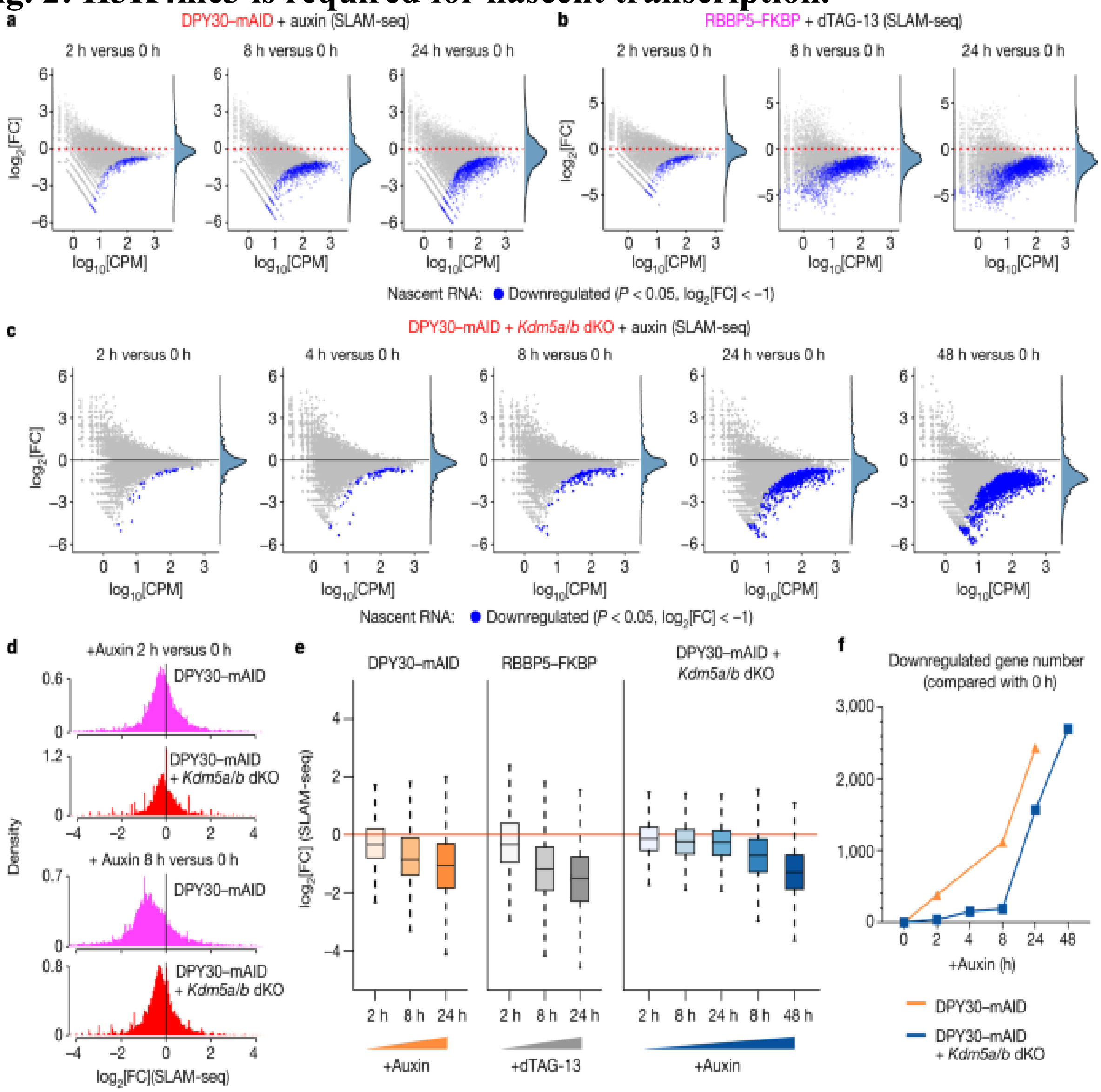
whether the rapid turnover of H3K4me3 was due to the activities of the KDM5 H3K4me3/me2 demethylases[22], we deleted *Kdm5a* and *Kdm5b* (*Kdm5a/b* double knockout (dKO)) in the DPY30–mAID mES cell line (DPY30–mAID *Kdm5a/b* dKO) (Fig. 1f). The deletion of the two genes did not lead to significant changes in cell proliferation and expression of pluripotent genes in the two independent isolated clones (Extended Data Fig. 2d,e). The dKO cells led to a global increase in H3K4me3 levels at steady state (Fig. 1g), consistent with previously reported data[22,23]. In contrast to in DPY30–mAID cells, in which global H3K4me3 was lost within 2 h, global H3K4me3 persisted for 8 h in dKO cells. By contrast, the turnover patterns of H3K4me1 and H3K4me2 were largely unaffected in the dKO cells (Fig. 1h). Furthermore, the *Kdm5a/b*-dKO cells showed a delayed reduction in H3K4me3 genome-wide relative to DPY30–mAID cells (Fig. 1e,i). Moreover, we found that the binding of the catalytic and non-catalytic components of the SET/COMPASS complexes was not significantly decreased after degradation of DPY30 (Extended Data Fig. 2f). These data demonstrate that the rapid turnover of H3K4me3 is dependent on KDM5A and KDM5B in mES cells.

H3K4me3 is required for transcription

To determine the primary effects on transcription after rapid H3K4me3 removal, we measured the synthesis of newly transcribed RNAs using thiol(SH)-linked alkylation for the metabolic sequencing of RNA (SLAM-seq)[24] in the degron cell lines (Extended Data Fig. 3a–d). Short-term treatment with auxin and dTAG-13 (2 h and 8 h) led to a significant reduction in mRNA synthesis (Fig. 2a,b). At these time points, H3K4me3 was lost, whereas there was a minimal effect on H3K4me2 and H3K4me1 levels (Fig. 1b–e). The number of downregulated genes ($P < 0.05$ and $\log_2$-transformed fold change < -1), as well as their magnitude of decreased expression, increased over time in both degron systems (Fig. 2a,b). Notably, the earliest downregulated genes were more likely to be regulated by a CGI-rich promoter, to show higher levels and broader peaks of H3K4me3 than genes with unaltered expression, and to be involved in processes such as ribosome biogenesis, translation and cell cycle progression (Extended Data Fig. 3e–h). Taken together, as only H3K4me3 (and not H3K4me2/me1) was

affected at early timepoints after acute loss of DPY30 or RBBP5, these data suggest that H3K4me3 is required for transcription.

Fig. 2: H3K4me3 is required for nascent transcription.



a, *MA* plots depicting changes in nascent transcription (SLAM-seq) at the indicated times after auxin treatment in DPY30–mAID cells. $n = 3$ biological replicates. CPM, counts per million mapped reads; FC, fold change. Adjusted *P* values were calculated using Wald tests in DESeq2. **b**, *MA* plots depicting changes in nascent transcription (SLAM-seq) at the indicated times after dTAG-13 treatment in RBBP5–FKBP cells. $n = 3$ biological replicates. Adjusted *P* values were calculated using Wald tests in DESeq2. **c**,

MA plots depicting changes in nascent transcription (SLAM-seq) at the indicated times after auxin treatment in DPY30–mAID *Kdm5a/b*-dKO cells. $n = 2$ biological replicates. Adjusted *P* values were calculated using Wald tests in DESeq2. **d**, The $\log_2$-transformed fold change in nascent gene expression in the depicted cell lines on the basis of the data shown in **a** and **c**. **e**, The nascent transcriptional changes ($\log_2$-transformed) for genes in indicated samples across timepoints with DPY30 or RBBP5 degradation kinetics. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × interquartile range (IQR) above and below the box (whiskers). $n = 3$ (DPY30–mAID or RBBP5–FKBP degron cells) and $n = 2$ (dKO cells) biological replicates. **f**, Comparison of the number of downregulated genes after auxin treatment for the indicated cell lines, based on the data presented in **a** and **c**.

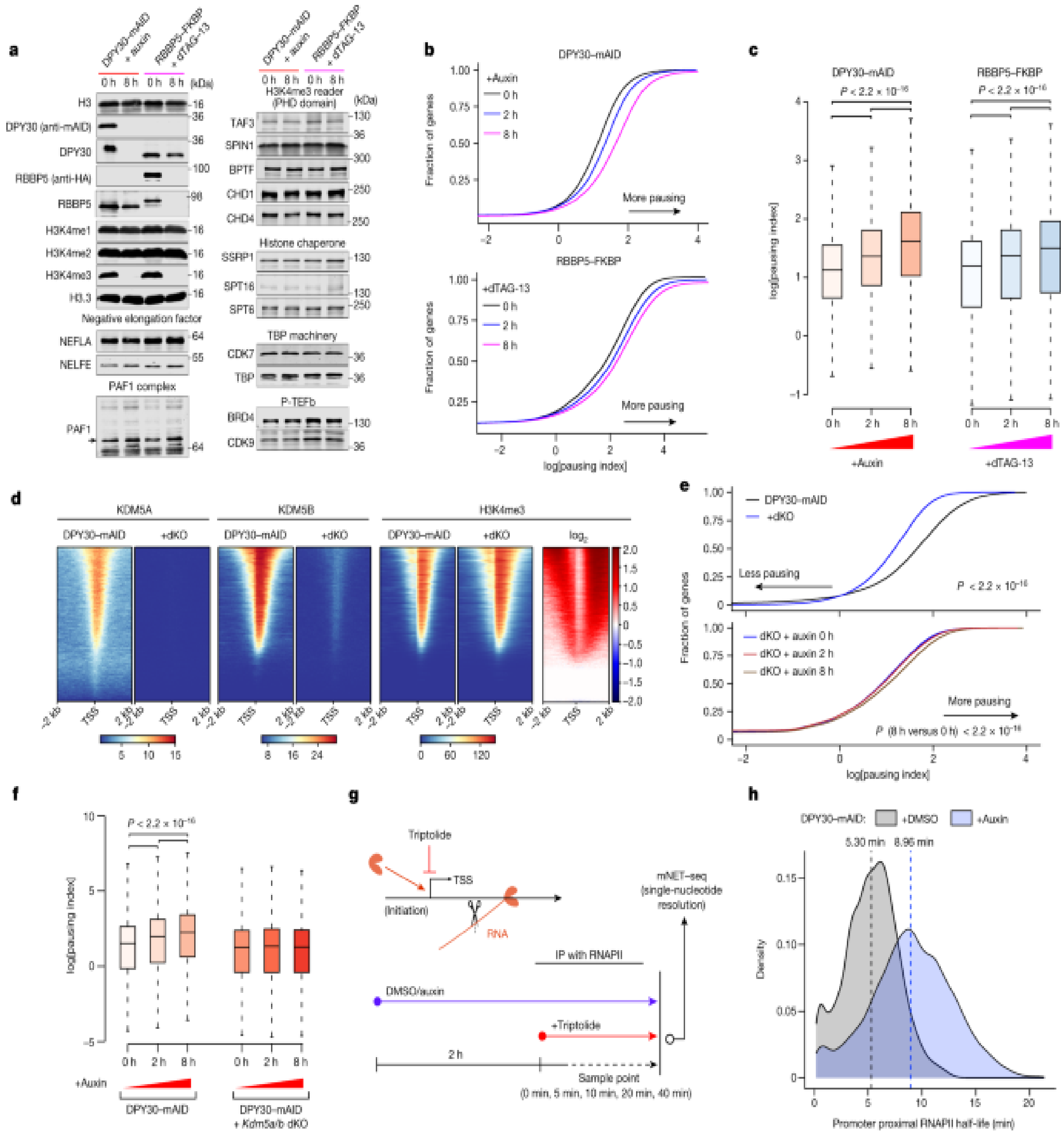
If H3K4me3 were the determining factor regulating transcription, deletion of *Kdm5a* and *Kdm5b* should also lead to a significant delay in gene expression changes in response to DPY30 loss. To test this, we performed SLAM-seq analysis of auxin-treated dKO cells (Fig. 2c). Indeed, we observed a significant delay in the decrease in mRNA synthesis in the dKO cells compared with in the DPY30–mAID cells (Fig. 2d,e). Specifically, only 41 and 186 genes were downregulated in the dKO cells 2 h and 8 h after auxin treatment, respectively, whereas 379 and 1,115 downregulated genes were found at the same timepoints in DPY30–mAID cells after auxin treatment, respectively (Fig. 2f). Thus, these results suggest that the reduction in H3K4me3 levels contributes to the observed decrease in transcription.

Intact PIC formation after loss of H3K4me3

As H3K4me3 has been reported to recruit proteins to enhance transcription initiation[1,25], we examined whether the loss of H3K4me3 could lead to changes in the expression of proteins in the RNAPII pre-initiation complex (PIC). As demonstrated in Fig. 3a, acute loss of H3K4me3 did not affect the global protein levels of multiple subunits of the PIC, such as CDK7 (TFIIH) and TBP (TFIID). Furthermore, reported H3K4me3-binding proteins also showed comparable protein levels and genome-wide binding patterns in control and auxin/dTAG-13-treated cells (Fig. 3a and Extended Data Fig. 4a). To address whether H3K4me3 loss would lead to changes in RNAPII

PIC formation, we affinity-purified proteins that associate with RNAPII and determined their identities using mass spectrometry (MS) analysis of samples in which DPY30 was depleted or not (Extended Data Fig. 4b,c). However, loss of DPY30 did not lead to significant changes in the expression of RNAPII-associated proteins or in their interaction with RNAPII (Extended Data Fig. 4d and Supplementary Table 3). Thus, the loss of H3K4me3 does not lead to detectable changes in the formation of the RNAPII PIC or the association with PIC subunits to TSSs.

Fig. 3: Acute loss of H3K4me3 increases the residence time of paused RNAPII.



a, Immunoblot analysis of the indicated transcriptional core proteins and H3K4me3 readers in the indicated cell lines treated with or without auxin or dTAG-13 as shown. **b**, The RNAPII pausing index in control (0 h, black) and auxin-treated or dTAG-13-treated degron cells. Higher index values indicate a higher degree of RNAPII pausing. Cumulative index plots of the pausing index were calculated from total RNAPII ChIP–seq signals. **c**, The RNAPII pausing index was determined using ChIP–seq in the indicated samples with DPY30 or RBBP5 degradation kinetics. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 ×

IQR above and below the box (whiskers). *P* values were calculated using two-sided Wilcoxon rank-sum tests. $n = 12{,}621$ genes. **d**, Comparison of the occupancy of KDM5A, KDM5B and H3K4me3 around the TSS region (TSS ± 2 kb) in DPY30–mAID and DPY30–mAID *Kdm5a/b*-dKO cells. **e**, The RNAPII pausing index in DPY30–mAID (black), DPY30–mAID *Kdm5a/b*-dKO (blue) and auxin-treated cells. Higher index values indicate a higher degree of RNAPII pausing on promoter region of genes. *P* values were calculated using two-sided Wilcoxon tests. **f**, The RNAPII pausing index determined using mNET–seq in DPY30–mAID, DPY30–mAID *Kdm5a/b*-dKO and auxin-treated cells. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × IQR above and below the box (whiskers). *P* values were calculated using two-sided Wilcoxon tests. $n = 10{,}332$ genes. **g**, The experimental strategy of the mNET–seq approach to measure the promoter-proximal RNAPII half-life after treatment with triptolide. **h**, Density plot showing increased paused RNAPII half-life of $n = 4{,}007$ genes after acute loss of H3K4me3. The average of paused RNAPII half-life is shown as a dashed line. $n = 2$ biological replicates.

H3K4me3 regulates RNAPII occupancy

To determine the relationship between H3K4me3 and RNAPII occupancy genome-wide, we performed a ChIP–seq analysis of total RNAPII in the degron cells. Consistent with the lack of detectable effects on RNAPII PIC formation, we did not observe decreased RNAPII enrichments at promoter regions after auxin or dTAG-13 treatment. By contrast, we found an increase in RNAPII occupancy at promoter regions (Extended Data Fig. 4e), which indicated that acute loss of H3K4me3 promoted RNAPII pausing. This was confirmed by calculating the RNAPII pausing index[26] in both DPY30–mAID and RBBP5–FKBP cells (Fig. 3b,c and Extended Data Fig. 4f,g). Consistent with the increased pausing, we observed an accumulation of the RNAPII CTD Ser5 phosphorylation (Ser 5p) and negative elongation factor A (NELFA) at promoter-proximal regions and a marked reduction in RNAPII CTD Ser2 phosphorylation (Ser 2p) throughout gene bodies (Extended Data Fig. 4h,i). These data suggest that H3K4me3 is involved in the regulation of RNAPII pausing.

To gain further support that the observed effect on RNAPII pause-release was caused by loss of H3K4me3, we determined the effect of degrading DPY30 in *Kdm5a/b*-dKO cells. We confirmed that KDM5A and KDM5B were lost from TSSs in the dKO cells using ChIP–seq, and that the loss of the KDM5s led to a corresponding increase in H3K4me3 levels under steady-state conditions (Fig. 3d). Moreover, we found that the increase in H3K4me3 corresponds to a concomitant decrease in steady-state RNAPII pausing in dKO cells (Fig. 3e and Extended Data Fig. 5a). We also observed a significant delay in the onset of RNAPII pausing in the dKO cells compared with in DPY30–mAID cells expressing KDM5A/B (Fig. 3e and Extended Data Fig. 5a–c). Moreover, we performed nascent transcription with mammalian native elongating transcript sequencing (mNET–seq)[27] at single-nucleotide resolution in the DPY30–mAID and dKO lines. Consistently, the mNET–seq data showed that loss of KDM5A/B led to a significant delay in RNAPII pausing compared with in KDM5A/B-expressing cells (Fig. 3f). Taken together, these results show a role for H3K4me3 in regulating RNAPII pause-release.

H3K4me3 regulates RNAPII half-life

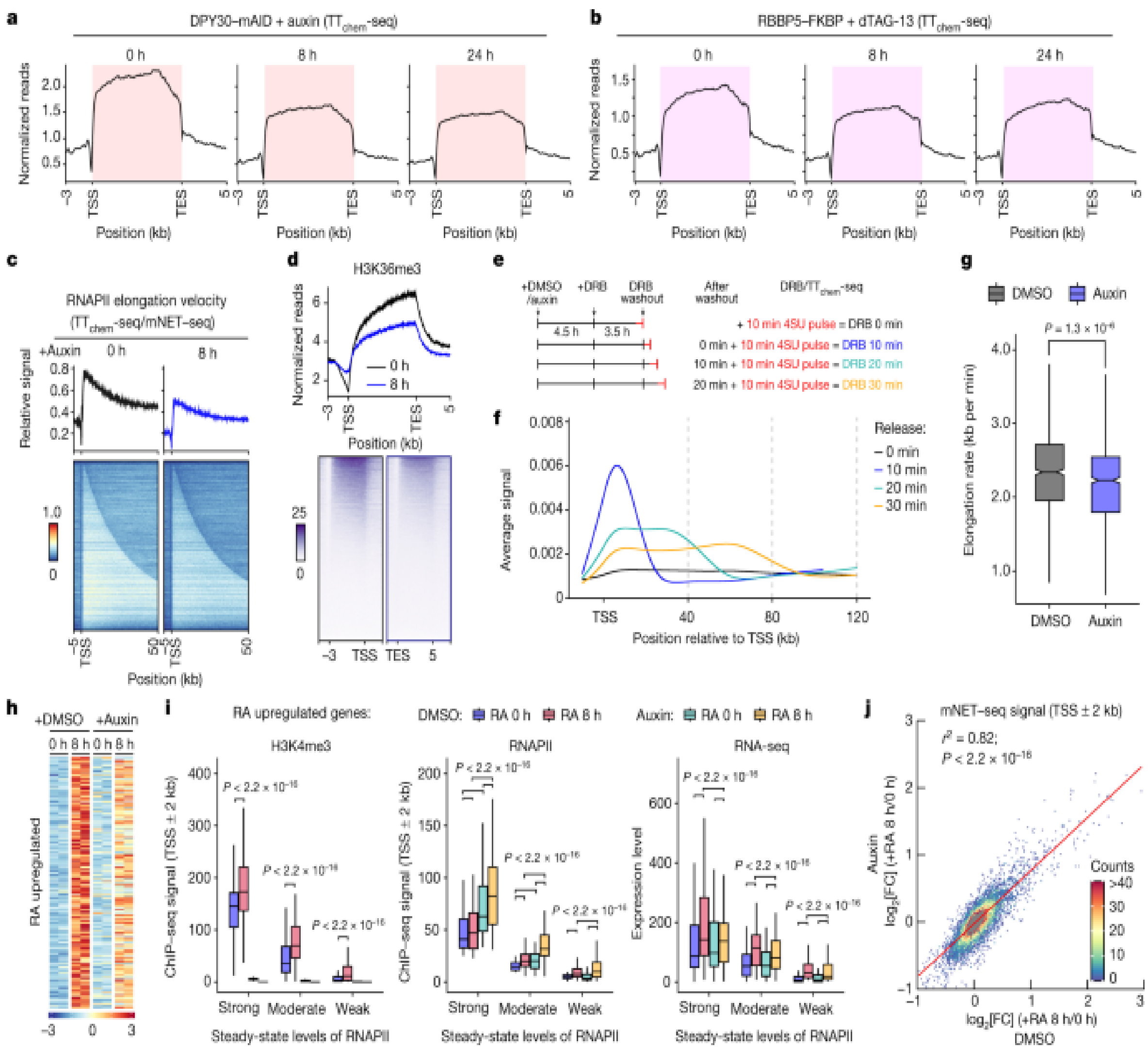
An increase in the occupancy of promoter-proximal RNAPII can be attributed to increased RNAPII initiation, blockage of RNAPII entering the gene body or both. To determine the stability and half-life of RNAPII at promoter-proximal regions, we inhibited transcription initiation with triptolide[28,29,30] and combined it with mNET–seq (Fig. 3g). The single-nucleotide resolution of the mNET–seq data enabled us to specifically measure the dynamics of RNAPII-engaged RNA at promoter-proximal regions. We fitted the RNAPII time-course measurements to an exponential decay model[28] and calculated the half-life of promoter-paused RNAPII on all protein-coding genes (Extended Data Fig. 6a). This analysis showed that the RNAPII half-life is 5.3 min at protein-coding genes in steady-state mES cells ($n = 4{,}007$ genes) and increased to 8.96 min ($P < 2.2 \times 10^{-16}$) in auxin-treated cells (Fig. 3h and Extended Data Fig. 6b,c). This change in RNAPII half-life is similar to what has been observed after CDK9 inhibition[31]. Indeed, we found that CDK9, BRD4 and HEXIM1 occupancies were increased at the promoter regions (Extended Data Fig. 4i) in both degron cell

lines. Thus, our data suggest that H3K4me3 regulates transcription by facilitating the release of paused RNAPII into productive elongation in mES cells.

H3K4me3 regulates elongation

To examine and monitor the effects of H3K4me3 on RNA synthesis rates genome-wide from actively transcribing RNAPII, we used a modified transient transcriptome sequencing (TT_{chem}-seq) method[32] with spike-in controls. Loss of H3K4me3 in auxin-treated or dTAG-13-treated cells led to a reduction of approximately 50% in the transcription of protein-coding genes at the 8 h timepoint (Fig. 4a,b) but resulted in increased RNAPII-engaged RNA at the promoter regions as measured by mNET–seq (Extended Data Fig. 6d). Furthermore, we found that there were no significant changes in RNAPII occupancy at either active or inactive enhancer regions after H3K4me3 loss (Extended Data Fig. 6e,f). To estimate the change in elongation velocity, we used the ratio of nascent RNA synthesis measured using TT_{chem}-seq and RNAPII RNA occupancy measured using mNET–seq (TT_{chem}-seq/mNET–seq) as a proxy for elongation velocity, which is a measurement of the amount of ongoing RNA synthesis per RNAPII molecule[33,34]. This analysis showed that acute loss of H3K4me3 caused a general transcriptome-wide decrease in elongation velocity (Fig. 4c). Consistent with this observation, the well-known epigenetic mark for transcription elongation H3K36me3 also showed a global decrease in the auxin-treated cells (Fig. 4d and Extended Data Fig. 6g).

Fig. 4: H3K4me3 regulates transcriptional elongation.



a,**b**, Metagene profiles for transient transcriptome sequencing (TT_{chem}-seq) in control and auxin-treated (**a**) or dTAG-13-treated (**b**) cells in the indicated cell lines. TES, transcription end site. **c**, Heat maps and profiles showing changes in elongation velocities (TT_{chem}-seq/mNET–seq) after acute loss of H3K4me3. **d**, H3K36me3 ChIP–seq profiles and heat maps in control and auxin-treated DPY30–mAID cells. **e**, Outline of the DRB/TT_{chem}-seq experiment to measure RNAPII elongation rates. 4SU, 4-thiouridine. DRB 0 min, no release of DRB. **f**, DRB/TT_{chem}-seq metagene profiles of protein-coding genes (60–300 kb length) with non-overlapping transcriptional units ($n = 3,566$) in the depicted cells. Lines are computationally fitted splines. **g**, Box plot showing decreased RNAPII elongation rates after H3K4me3 loss. P values were calculated using two-sided Wilcoxon tests. $n = 855$ genes with

RPM > 100. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × IQR above and below the box (whiskers). **h**, The upregulated genes response to RA treatment in auxin-treated and DMSO-treated cells ($n = 2$). Gene expression is shown as relative *Z*-scores across the samples. **i**, The changes in H3K4me3 and RNAPII ChIP–seq at TSSs (±2 kb), and RNA-sequencing analysis of RA-response genes (upregulated genes) in the indicated samples. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × IQR above and below the box (whiskers). *P* values were calculated using two-sided Wilcoxon tests. $n = 77$ genes for each group. **j**, The correlation of mNET–seq signal around the TSS region (TSS ± 2 kb) at 8 h after RA treatment with or without H3K4me3. Pearson correlation and *P* values are reported at the top. *P* values were calculated using two-sided Wilcoxon tests.

As an orthogonal assay to investigate how H3K4me3 loss affects transcription elongation, we determined productive RNAPII elongation rates using TT_{chem}-seq in combination with the reversible CDK9 inhibitor 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) (DRB/TT_{chem}-seq)[32]. The progression of RNAPII was reflected by metagene coverage profiles after DRB release and clear transcription wave peaks on genes (Fig. 4e,f). The average elongation rates for the 3,655 non-overlapping protein-coding genes (60 kb to 300 kb in length) were 2.2 kb per min in control cells (Extended Data Fig. 6h), which is in good agreement with data obtained in HEK293[32] and HCT116[35] cells. We found that loss of H3K4me3 led to a significant decrease in RNAPII elongation rates (Fig. 4g and Extended Data Fig. 6i,j). Taken together, these observations suggest that H3K4me3 regulates both RNAPII pause-release and elongation in mES cells.

Role of H3K4me3 in initiation

It has previously been suggested, mainly on the basis of in vitro experiments[1,25], that H3K4me3 functions to facilitate the formation of the PIC at TSSs. To test this hypothesis, we determined whether H3K4me3 is required for de novo activation of transcription in response to retinoic acid (RA)-induced differentiation (Extended Data Fig. 7a). A principal component analysis and correlation analysis of the entire transcriptome

showed that loss of H3K4me3 did not impair the overall rewiring of gene expression induced by RA treatment (Fig. 4h and Extended Data Fig. 7b–e), suggesting that H3K4me3 is dispensable for the transcription initiation of these genes.

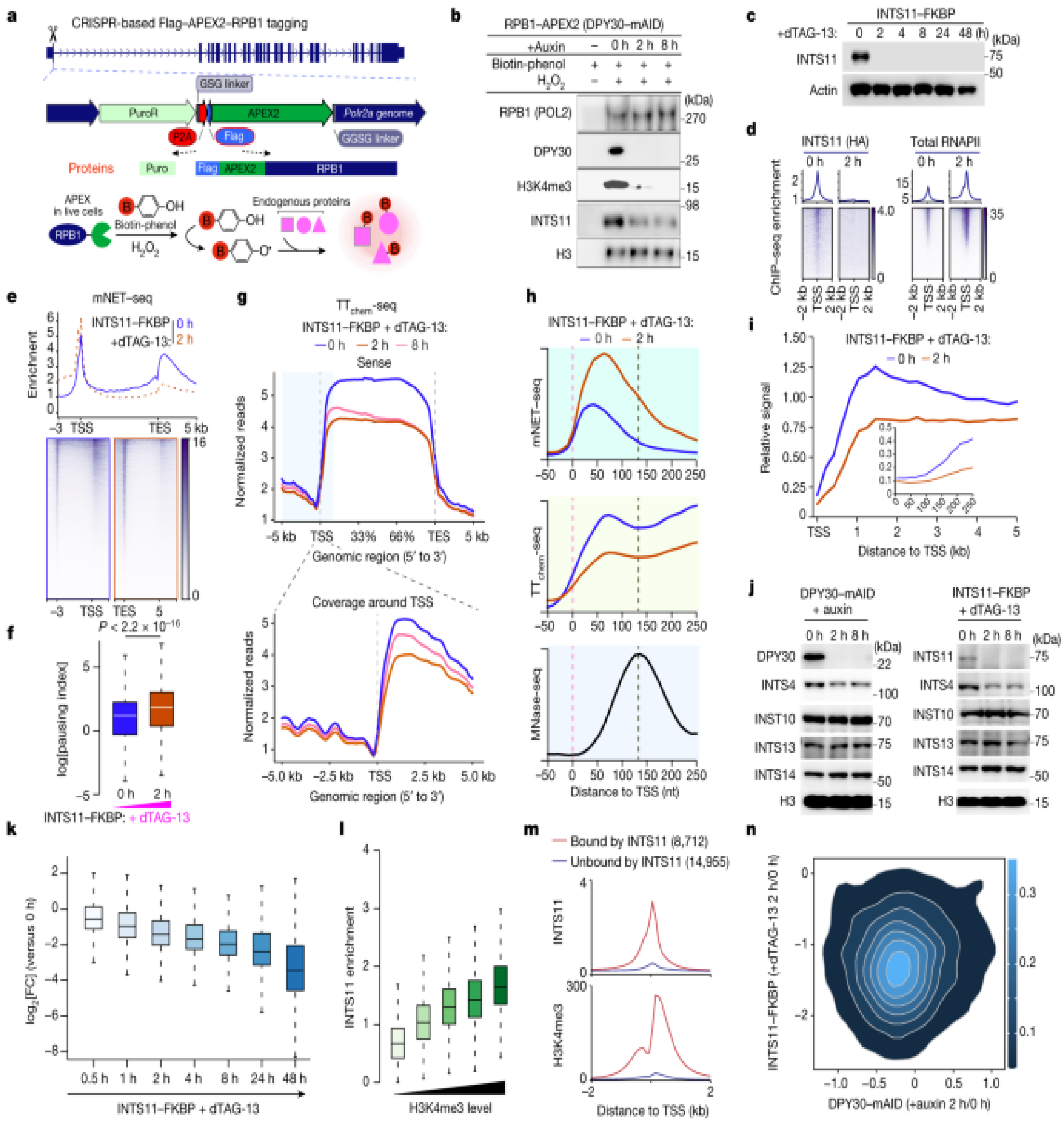
Considering that many RA-response genes are bivalent with both low levels of gene expression and the histone modifications H3K4me3 and H3K27me3[36] at steady state, they may already be primed for de novo transcription. To analyse this in more detail, we determined H3K4me3 and RNAPII location using ChIP–seq analysis of DPY30–mAID cells after RA treatment with or without auxin treatment. As expected, H3K4me3 was lost in auxin-treated cells and showed increased RNAPII at promoter regions (Extended Data Fig. 6f,g). By dividing the RA-induced genes into three categories on the basis of the levels of RNAPII enrichment (strong, moderate and weak) at promoter regions in steady-state mES cells (Fig. 4i), we found increased RNAPII enrichments and gene expression for the same three category genes in the auxin-treated cells (Fig. 4i), suggesting that the RA-response genes can be initiated de novo in the absence of H3K4me3. As these steady-state data cannot rule out an effect on transcriptional initiation, high-resolution mNET–seq was performed. Correlation analysis of the mNET–seq results also showed that the loss of H3K4me3 did not impair the loading of RNAPII at the promoter-proximal region of genes induced by RA treatment (Fig. 4j). Thus, we conclude that H3K4me3 is not required for RNAPII loading and for transcriptional initiation.

The H3K4me3-dependent RNAPII interactome

To understand the mechanism leading to increased promoter-proximal pausing in response to H3K4me3 loss, we combined CRISPR-based genome editing, APEX2-based proximity labelling[37] and quantitative MS (SILAC/MS) to obtain a high-resolution view of the molecular and spatial organization of RNAPII with or without H3K4me3 (Fig. 5a). We knocked-in APEX2–Flag-tagged *Rbp1* in both degron cell lines, which did not lead to detectable effects on the cellular functions of the protein (Extended Data Fig. 8a–e). After activation with H_2O_2, APEX2 oxidizes phenol derivatives (biotin-phenol), which covalently react with the nearby endogenous proteins (Extended Data Fig. 8f,g). To capture the changes on chromatin after the

loss of H3K4me3, we isolated the chromatin fraction from each sample before mixing the light and heavy conditions for MS (Extended Data Fig. 8h). We identified 1,901 proteins that were in the proximity of RPB1, and KEGG and Gene Ontology analyses showed that most of these proteins are associated with RNA biogenesis processes (Extended Data Fig. 8i–l and Supplementary Table 4), highlighting the quality of the RPB1–APEX2 data.

Fig. 5: INTS11 regulates pause-release and transcription dependent on H3K4me3.



a, The strategy for CRISPR-based Flag–APEX2–RPB1 (RNAPII–APEX2) tagging. **b**, Validation of H3K4me3-dependent INTS11 chromatin interaction in DPY30–mAID cells. **c**, Western blot analysis of HA-tagged INTS11 and actin in INTS11–FKBP cells. **d**, HA-tagged INTS11 and total RNAPII ChIP–seq profiles and heat maps in INTS11–FKBP cells. **e**, mNET–seq profiles and heat maps in INTS11–FKBP degron cells. **f**, The RNAPII pausing index was determined using mNET–seq in INTS11–FKBP cells. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × IQR above and below the box (whiskers). *P* values were calculated using two-sided Wilcoxon tests. *n* = 10,332 genes. **g**, Metagene transcriptional profiles were acquired using TT_{chem}-seq in INTS11–FKBP degron cells. **h**, Metagene analyses of mNET–seq and TT_{chem}-seq signals at single-nucleotide resolution acquired in INTS11–FKBP cells. MNase-seq, micrococcal nuclease sequencing. **i**, Analysis of elongation velocities (TT_{chem}-seq/mNET–seq) after acute loss of INTS11. **j**, Immunoblot analysis of integrator subunits in the indicated cell lines treated with or without auxin or dTAG-13 as shown. **k**, Box plot comparing the log_2-transformed fold change in SLAM-seq at the indicated timepoints. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × IQR above and below the box (whiskers). *n* = 13,776 genes. **l**, The INTS11 signal at TSSs (± 2 kb) in the indicated groups. The box plots indicate the median (centre line), the third and first quartiles (box limits) and 1.5 × IQR above and below the box (whiskers). *n* = 4,700 genes. **m**, The average distribution of INTS11 and H3K4me3 ChIP–seq signals at INTS11-bound genes (*n* = 8,712) versus INTS11-unbound genes (*n* = 14,955) in mES cells. **n**, 2D kernel density plot showing the relationship between SLAM-seq changes in INTS11–FKBP and DPY30–mAID degron cells. The colour bar reflects the intensity.

We focused on the changes in the two early timepoints (2 h and 8 h) for which only H3K4me3 (and not H3K4me2 and H3K4me1) was lost in response to auxin treatment (Extended Data Fig. 8m). Importantly, a large proportion of the proteins that showed differential interactions with RPB1 after DPY30 degradation were shared at the two times (Extended Data Fig. 8n). Consistent with the experimental design, DPY30 was the top downregulated protein in all of the auxin-treated samples (Extended Data Fig. 8o,p). Moreover, functional enrichment analysis of the common 228

downregulated proteins indicated a strong enrichment for mRNA-processing components of the core RNAPII machinery (Extended Data Fig. 8q). In addition to H3K4me3 loss in the cells, terms such as 'positive regulation of histone H3K4 methylation', 'chromatin binding' and 'histone binding' were more prominently decreased with chromatin RNAPII (Extended Data Fig. 8q).

H3K4me3-dependent recruitment of INTS11

To identify proteins that could potentially explain the requirement for H3K4me3 in promoter-proximal pause-release, we overlapped the common downregulated proteins with proteins that have previously been shown to associate with H3K4me3 using cross-linked ChIP–MS[38]. Notably, three proteins were in common between the two protein groups: DPY30 itself, PAF1 and INTS11 (Extended Data Fig. 9a). These three proteins were also found to be preferentially enriched in the H3K4me3 ChIP–MS data, when compared with ChIP–MS data from other heterochromatin or non-promoter specific histone modifications (Extended Data Fig. 9b). PAF1[39,40] and INTS11[41,42] have both been reported to have important roles in RNAPII pause-release and also in transcriptional elongation. We chose to focus on INTS11—an endonuclease subunit of the Integrator complex—as it showed a more substantial decrease in RPB1 association than PAF1 in response to H3K4me3 depletion (Extended Data Fig. 8p).

The reduction in RNAPII-bound INTS11 after acute loss of H3K4me3 was further confirmed by western blot analysis of both DPY30–mAID and RBBP5–FKBP degron cells (Fig. 5b and Extended Data Fig. 9c). Recent studies indicate that integrator primarily attenuates gene expression in *Drosophila*[43] and in human HeLa[44] cells, while other studies have shown a critical role of INTS11 in gene activation and elongation[45,46]. To improve our understanding of the role of INTS11 in transcription regulation, we generated INTS11–FKBP-knockin cells along with a double haemagglutinin (HA) epitope tag in the DPY30–mAID RPB1–APEX2 degron mouse ES cell line (Extended Data Fig. 9d). Western blot analysis detected similar levels of INTS11 in the constructed cell line and in non-tagged cells, and the tagging of INTS11 did not lead to detectable effects on the expression of pluripotency genes and cell proliferation (Extended Data Fig. 9e–g). The

addition of dTAG-13 to these cells led to rapid degradation of the INTS11–FKBP–HA fusion protein within 2 h (Fig. 5c). Consistent with our findings for the role of H3K4me3 in regulating RNAPII pause-release, acute loss of INTS11 led to increased RNAPII pausing at promoter-proximal regions of genes (Fig. 5d–f) and to inhibition of cell proliferation (Extended Data Fig. 9g). Moreover, the enrichment of INTS11 at TSSs was significantly reduced in response to DPY30 degradation (Extended Data Fig. 9h). These data suggest H3K4me3-dependent recruitment of INTS11, which is required to enable transcriptional pause-release.

We further studied the effects of INTS11 loss by performing TT_{chem}-seq analysis of INTS11 degron cells (Extended Data Fig. 9i). Loss of INTS11 led to a significant increase in the RNA synthesis and RNAPII of non-coding short transcripts, such as upstream antisense RNAs (uaRNAs/PROMPTs) (Extended Data Fig. 9j,k). By contrast, a global decrease in the expression of protein-coding genes (Fig. 5g), especially at the pausing sites between TSS and the first (+1) nucleosome (Fig. 5h) was observed after INTS11 degradation. Analysis of elongation velocity at protein-coding genes showed broadly decreased productive elongation after INTS11 loss (Fig. 5i). Notably, the degradation of INTS11 led to the loss of only part of the integrator complex from chromatin (Fig. 5j). Taken together, these findings further support a functional link between H3K4me3 and INTS11.

To further investigate this link, we measured ongoing transcription using SLAM-seq analysis of the INTS11–FKBP degron cell line (Extended Data Fig. 10a). Notably, short-term treatment with dTAG-13 led to a significant decrease in nascent mRNA synthesis (Fig. 5k and Extended Data Fig. 10b). We also observed a strong correlation between the enrichments of H3K4me3 and INTS11 at TSS promoter-proximal regions (Fig. 5l) and that there was a significant enrichment of H3K4me3 at INTS11-bound genes, but not at INTS11-unbound genes (Fig. 5m and Extended Data Fig. 10c,d). By plotting the fold changes in nascent RNA expression determined in DPY30–mAID cells in response to auxin treatment and in INTS11–FKBP cells treated with dTAG-13, we also showed that most of the changes were correlated between the two degron systems (Fig. 5n).

Finally, we investigated whether similar correlations exist in other cell types and analysed published data from different human cell lines—HeLa[42], THP1[47] and HL60[48] (Supplementary Fig. 1a). ChIP–seq experiments in HeLa, THP1 and HL60 cells identified that ~60%, ~52% and ~61% of INTS11 peaks overlapped with promoters, respectively (Supplementary Fig. 1b). Notably, compared with the H3K4me1-occupied regions, most of the H3K4me3-occupied regions were marked with significant INTS11 and RNAPII levels (Supplementary Fig. 2a). In support of the results obtained in mES cells, INTS11 occupancy was heavily enriched towards the promoter-proximal region of genes in all of these human cell lines (Supplementary Fig. 2b). Moreover, heat maps also demonstrated a significant colocalization of the peak binding sites for INTS11 and RNAPII with the peak sites of H3K4me3 enrichment on active promoters (Supplementary Fig. 2b). Thus, these published ChIP–seq results from different cells validated a correlation between INTS11-binding sites and sites of H3K4me3 enrichment across the genome. Taken together, these results show that H3K4me3 regulates promotor-proximal pausing through a mechanism involving the recruitment of the INTS11, which is essential for the eviction of paused RNAPII and transcriptional elongation.

Discussion

H3K4me3, which is catalysed by the SET1/COMPASS complexes, is tightly associated with TSSs and is widely believed to be involved in regulating transcription initiation[1,2,7,25]. By generating and using model systems to study the direct role of H3K4me3 in regulating transcription, we have shown that loss of H3K4me3 leads to an increase in RNAPII at promoters. Moreover, we have shown, through multiple orthogonal assays, that H3K4me3 regulates RNAPII pausing and potentially also has a role in elongation. A connection between H3K4me3 and elongation has previously been suggested on the basis of results in cancer cells[12] and plants[49]. Notably, we did not detect a role for H3K4me3 in transcriptional initiation and activation of de novo transcribed genes in response to RA-induced differentiation. Furthermore, we observed a fast turnover of H3K4me3 that is dependent on KDM5 demethylase activity, and we propose that this rapid turnover of H3K4me3 is important for the dynamic regulation of transcriptional output. Our experimental approach does not enable us to

delineate specific roles for H3K4me1 and H3K4me2 in transcriptional regulation because all H3K4 methylation states are influenced at later time points. Notably, recent results suggest that SET1A/B can also regulate CpG-island-associated gene expression independently of H3K4me3 and its methyltransferase activity[50], indicating potential distinct roles of components of the SET1/COMPASS complexes.

Paused RNAPII has been proposed to be coupled with transcription and mRNA-processing events by helping to maintain the accessibility of promoters for regulatory factors that activate transcription by recruiting P-TEFb and other factors, and promoting the release of paused RNAPII into gene bodies[30,33,39]. Whether histone modifications have a role in regulating the RNAPII pause-release step is not well understood. Our results show that H3K4me3 coupled with INTS11 is involved in regulating RNAPII promoter-proximal pausing. We propose a model in which H3K4me3 regulates transcriptional cycles by facilitating the recruitment of INTS11 to protein-coding genes (Extended Data Fig. 11). INTS11 can subsequently, through its RNA endonuclease activity, mediate productive transcriptional elongation by evicting paused RNAPII. This model is supported by recent studies showing that mammalian INTS11 facilitates RNAPII pause-release and gene expression[45,51]. Most of our conclusions are also supported by another recent study; however, in this study the authors observed an upregulation of protein-coding short transcripts after INTS11 downregulation[52]. Although we do not know the reason for the different observations, it may reflect the efficiency of the two different degron systems used in the studies.

Further work will be needed to determine the detailed molecular mechanism that leads to the recruitment of INTS11 and its function in elongation. In summary, our study demonstrates that H3K4me3 is indeed an important post-translational modification that regulates promoter-proximal pause-release and facilitates gene expression.

Methods

mES cell culture and differentiation

E14 ES cells (129/Ola background) were maintained on 0.2% gelatin-coated plates in Glasgow minimum essential medium (GMEM, Sigma-Aldrich, G5154) containing 15% fetal bovine serum (Gibco, 26140079), supplemented with 1× penicillin–streptomycin (Thermo Fisher Scientific, 15140122), 2 mM GlutaMax (Thermo Fisher Scientific, 35050061), 50 µM β-mercaptoethanol (Thermo Fisher Scientific, 21985023), 0.1 mM non-essential amino acids (Thermo Fisher Scientific, 11140050), 1 mM sodium pyruvate (Thermo Fisher Scientific, 11360070) and Leukaemia inhibitory factor (LIF, 1,000 U ml^{-1}, Millipore), referred to as serum mES cell medium. Cells were passaged every 2 days by aspirating the medium, dissociating the cells with trypsin/EDTA solution (TE) briefly at room temperature before rinsing and dissociation in mES cell medium by pipetting. Cells were pelleted by centrifugation at 300*g* for 5 min. mES cell transfection was performed using Lipofectamine 3000 (Thermo Fisher Scientific, L3000-001) according to the manufacturer's instructions. Cell counts were performed using the Countess II automated cell counter (Thermo Fisher Scientific, AMQAX1000) using 10 µl of cell suspension and 10 µl of Gibco Trypan Blue Solution (Gibco, 15250061) according to manufacturer's instructions. For all-*trans*-retinoic acid (RA, Sigma-Aldrich, R2625-50MG) treatment, cells were induced with 1 µM RA and without LIF for the indicated time. During differentiation, RA medium was changed every 24 h. For SILAC experiments, cells were cultured in SILAC DMEM (Thermo Fisher Scientific, A33822) containing 15% dialysed FBS (Thermo Fisher Scientific, 26400044), to which either $^{13}C_6{}^{15}N_2$ l-lysine-2HCl (Thermo Fisher Scientific, 88209) and $^{13}C_6{}^{15}N_4$ l-arginine-HCl (Thermo Fisher Scientific, 89990) (heavy), or l-lysine (Sigma-Aldrich, L8662) and l-arginine (Sigma-Aldrich, L8094) containing only light isotopes (light) was added. All cell lines were subjected to STR authentification through ATCC and were tested for mycoplasma contamination.

Generation of mES cell knockin cell lines

For the generation of the auxin-inducible degradation system for DPY30 sgRNA targeting the stop-codon region were cloned into eSpCas9(1.1)-T2A-eGFP. Left and right homology arms, as well as the mAID-T2A-BFP middle part were ligated into a modified pUC19 vector backbone (a gift from S. Pollard) using the In-Fusion cloning kit (Takara, 638910). mES cells

were co-transfected with sgRNA- and donor-vector using Lipofectamine 3000 and sorted 48 h later for GFP/BFP-double-positive cells. Homozygous clones were then transfected with pPB-hygro-OsTIR1-P2A-mCherry and pBase plasmids and selected with 100 μg ml^{-1} hygromycin B (Thermo Fisher Scientific, 10687010). For the generation of the endogenous dTAG-inducible degradation system for RBBP5, sgRNA targeting the stop codon region were co-transfected into the cells and contained the following elements: left and right homology arms, as well as FKBP12(F36V), 2× HA tags, P2A and a neomycin-resistance gene. The transfected cells were selected with 100 μg ml^{-1} Geneticin selective antibiotic (G418 Sulfate) (Thermo Fisher Scientific, 10131027), single-cell sorted to obtain clonal cell lines and screened for correct biallelic integration. All homozygous insertions and knock-ins were confirmed by Sanger sequencing and western blotting. A list of the oligos and the sequences of the sgRNAs is provided in Supplementary Table 1.

Generation of mES knockout cell lines

Cells were transfected with eSpCas9(1.1)-T2A-eGFP or eSpCas9(1.1)-T2A-mCherry vectors containing a sgRNA targeting the specific genomic locus, respectively, and cells were single-cell sorted 48 h after transfection. To generate the *Kdm5a/b*-dKO cells in the DPY30–mAID cell line, we first generated the *Kdm5a*-KO in the DPY30–mAID line (*Kdm5a*$^{-/-}$) by Cas9 (sgRNAs are listed in Supplementary Table 1), then we generated the *Kdm5b* knockout in the *Kdm5a*$^{-/-}$ line (at passage three of the *Kdm5a*$^{-/-}$ line). Subsequently, two clones with both *Kdm5a* and *Kdm5b* knockout (referred to as DPY30–mAID *Kdm5a/b*-dKO in this study) were picked up for the downstream analysis. All homozygous insertions and knockouts were confirmed by Sanger sequencing and western blotting. A list of the sgRNAs is provided in Supplementary Table 1. Further characterization of the dKO cell lines showed that they did not have detectable changes in proliferation and expressed normal levels of pluripotent genes; however, as expected, they showed an increase in H3K4me3 levels as compared with the wild-type cells.

Western blotting

Cells were lysed in RIPA buffer with Halt protease inhibitor (Thermo Fisher Scientific, 78429). Proteins that were separated by SDS–PAGE using acrylamide gels (BioRad gel system) were transferred onto nitrocellulose membranes (LI-COR, 926-31092). The membranes were blocked in 5% skimmed milk (Sigma-Aldrich) in PBS-T (0.1% Tween-20 in PBS) and incubated with the primary antibody of interest (Supplementary Table 2). As secondary antibodies, either IRDye 800CW goat anti-rabbit IgG (925-32211, LI-COR Bioscience, 1:15,000) or IRDye 800CW goat anti-mouse IgG (925-32210, LI-COR Bioscience, 1:15,000) was used. Proteins were imaged using Image Studio Lite (Odyssey CLx imager, Li-COR Biosciences). Immunoblotting source data are provided in Supplementary Fig. 4.

Flow cytometry

mES cells were dissociated with trypsin/EDTA, resuspended in culture medium, centrifuged and resuspended in PBS. For intracellular flow cytometry, 0.5 ml of cold fixation buffer (BioLegend, 420801) was added and then incubated at room temperature for 10 min. Subsequently, the cells were labelled with the unconjugated rabbit DPY30 antibodies (Bethyl Laboratories, A304-296A) and subsequently with a FITC-conjugated goat anti-rabbit IgG antibody. Flow cytometry was performed using the Beckman Coulter CytoFlex system. The gating strategy is shown for the E14 sample in Supplementary Fig. 5.

RNA extraction, cDNA synthesis and RT–qPCR analysis

Total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, 74134) according to the manufacturer's protocol. One microgram of total RNA was subjected to reverse transcription using Transcriptor Universal cDNA Master (Sigma-Aldrich, 5893151001). qPCR with reverse transcription (RT–qPCR) reactions were set up in triplicate using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific, A25778) and primers (listed in Supplementary Table 1). Relative quantitation was performed to a housekeeping gene using a $\Delta\Delta C_t$ method, as indicated in the corresponding figure legends. Statistical analysis was performed using GraphPad Prism v.7 (GraphPad).

3′-RNA Quant-seq and SLAM-seq

Cells (1×10^6 per treatment condition) were resuspended in 350 µl of buffer RLT plus, and total RNA was extracted from cell pellets using the RNeasy Plus Mini kit (Qiagen, 74134). For SLAM-seq experiments[24], cells were incubated with 100 µM 4-thiouridine (4SU; Biosynth, NT06186) for 60 min before RNA isolation. RNA (1 µg) was treated with 10 mM iodoacetamide in a 50 µl reaction volume at 50 °C for 15 min with 50 mM NaH_2PO_4 (pH 8.0), and 50% (v/v) DMSO followed by addition of 1 µl of 1 M dithiothreitol (DTT) to stop the reaction. RNA was precipitated at −80 °C for 60 min with 1 µl of GlycoBlue (Thermo Fisher Scientific, AM9515), 5 µl of 3 M sodium acetate (pH 5.2) and 180 µl of ethanol (≥99.0%). RNA was pelleted at 4 °C (12,000*g* for 30 min), washed with 1 ml of 80% ethanol and centrifuged at 4 °C (12,000*g* for 30 min). The RNA pellet was dried at room temperature for 10 min and resuspended in 20 µl of nuclease-free H_2O. RNA yield and quality were assessed using the 2200 TapeStation (Agilent). Sequencing libraries were prepared using the QuantSeq 3′-mRNA Seq Library Prep Kit FWD for Illumina (Lexogen, SKU: 015.96) from 500 ng or 100 ng of total RNA spiked with ERCC RNA Spike-In Mix 1 (1:1,000, Thermo Fisher Scientific, 4456740). In brief, first-strand (oligo(dT)) cDNA synthesis was followed by RNA removal and second-strand synthesis by random priming. The double-stranded library was bead-purified to remove reaction components before PCR amplification with i7 single-index primers for 10 cycles. Amplified libraries were again bead-purified according to the manufacturer’s protocol, and the concentration was measured by Qubit assay. All of the samples were checked for fragment size distribution on the TapeStation before pooling for 75 bp or 45 bp single-end read sequencing on the Illumina NextSeq 550 platform.

ChIP–seq

ChIP experiments were performed according to a standard protocol. In brief, ES cells were cross-linked by the addition of 1% formaldehyde (Sigma-Aldrich, 252549-1L) in the dish for 10 min at room temperature before quenching with 0.125 M glycine. The fixed cells were washed with PBS and resuspended in SDS buffer (100 mM NaCl, 50 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.5% SDS, 1× protease inhibitor cocktail from Roche). The resulting

nuclei were precipitated, resuspended in the immunoprecipitation buffer at 1 ml per 16 million cells (SDS buffer and Triton dilution buffer (100 mM NaCl, 100 mM Tris-HCl pH 8.0, 5 mM EDTA, 5% Triton X-100) mixed at a 2:1 ratio with the addition of 1× protease inhibitor cocktail from Roche) and processed on the Bioruptor Plus Sonicator (Diagenode) to achieve an average fragment length of 200–300 bp. Chromatin concentrations were estimated using the Nanodrop according to manufacturer's protocols. The immunoprecipitation reactions were set up in 1 ml of the immunoprecipitation buffer as indicated below and incubated overnight at 4 °C. The next day, BSA-blocked Protein G Dynabeads (Thermo Fisher Scientific, 10004D) were added to the reactions and incubated for 2 h at 4 °C. The beads were then washed three times with low-salt washing buffer (150 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA, 20 mM Tris-HCl pH 8.0) and twice with high-salt washing buffer (500 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA, 20 mM Tris-HCl pH 8.0). The samples were then reverse cross-linked overnight at 65 °C in the elution buffer (1% SDS, 0.1 M $NaHCO_3$) and purified using the QIAQuick PCR purification kit (Qiagen, 28506). A list of the antibodies used in this study is provided in Supplementary Table 2. Libraries for ChIP–seq were prepared using the NEBNext Ultra II DNA Library prep kit (NEB, E7645L), and AmpureXP beads (Beckman, A63881) were used for size selection. Libraries were quantified using the Qubit High Sensitivity DNA kit (Agilent, Q32854) and assessed on the TapeStation. Libraries were pooled as required, denatured and loaded onto the Illumina NextSeq 550 system with high-output kits (75 cycles). A list of all of the primers used for ChIP–qPCR is provided in Supplementary Table 1. For spiked-in ChIP–seq, 5% of the cross-linked *Drosophila* chromatin (homemade) with Spike-in Antibody (Active Motif, 61686) was added before the immunoprecipitation step according to the manufacturer's instructions.

RNAPII IP followed by MS

To measure the difference between RNAPII interactome in the presence and absence of DPY30, 1×10^8 DPY30–mAID cells were incubated with auxin for 8 h to induce DPY30 degradation (8 h samples), while 1×10^8 DPY30–mAID cells treated with DMSO served as the control (0 h samples). Cells were collected and frozen on dry ice and kept at −80 °C until

immunoprecipitation (IP). The cells were thawed at 37 °C for 30 s lysed in 1.6 ml of ice cold 50 mM EPPS pH 7.5, 150 mM NaCl, 1% Triton X-100 with cOmplete, EDTA-free Protease Inhibitor Cocktail (1 tablet per 20 ml of lysis buffer), 1:100 of Sigma-Aldrich phosphatase inhibitor 2 and 3 cocktails and 250 U μl^{-1} of benzonase. Lysates were incubated on ice for 5 min to allow DNA digestion, centrifuged at 20,000*g* for 5 min to remove insoluble material and filtered through the AcroPrep 1.0 μm glass filter plate at 2,000*g* for 1 min. The concentration of protein was then estimated using the bicinchoninic acid assay. For each of the samples (DMSO- and auxin-treated cells) six immunoprecipitation reactions were performed: three with the anti-RNAPII antibody (Abcam, ab817, 8WG16) and three with IgG control (Invitrogen, 02-6102). Each reaction was performed with 1 mg of lysate at 3.3 g l^{-1} lysate concentration and 5 μg of antibody bound to protein G Sepharose (Sigma-Aldrich, 17-0618-02). The incubation was performed at 4 °C with shaking at 1,100 rpm for 1 h. The beads were then transferred to the OF 1100 filter plate (Orochem Technologies) and washed five times with ice-cold 50 mM EPPS pH 7.5, 150 mM NaCl using vacuum manifold. Then, 18 μl of 10 mM EPPS pH 8.5 with 20 ng μl^{-1} trypsin and 1 ng μl^{-1} LysC was added to the beads in each well and digestion was performed for 2 h at 37 °C at 2,000 rpm. The partial digest was then collected into a 96-well PCR plate and left overnight at room temperature to complete digestion. Then, 4 μl of 22 g l^{-1} 11 plex TMTPro tags were added to each sample. The samples were then pulled and 20 μl of the combined sample was set aside, and the rest was fractionated into six fractions using the High pH Reversed-Phase Peptide Fractionation Kit, as suggested by the manufacturer. The fractions were concatenated into four fractions (the first and fifth fractions, the second and sixth and so on were mixed) and evaporated in speed vac (0.5 μl of DMSO was added to each sample to prevent complete evaporation) and resuspended in 20 μl 0.1% TFA. For data acquisition, 4.5 μl of unfractionated sample and every fraction was analysed by using the nanoAcquity 2 μm particle size, 75 mm × 500 mm easyspray column in direct injection mode. The samples were separated using the following gradient at 300 nl min^{-1} of buffer A (0.1% formic acid in water) and buffer B (0.1% formic acid in acetonitrile): 0–7% in 10 min, 7–30% in 92 min, 30–60% in 18 min, the column was then washed with 95% B for 10 min at 400 nl min^{-1}. The column was kept at 60 °C. Eluting peptides were analysed on the Orbitrap Fusion Lumos instrument using MS3 SPS with the settings recommended by the instrument

manufacturer for TMT11 plex analysis with the following modifications: (1) CID NCE for MS2 was set at 32; (2) HCD NCE for MS3 was set at 45; (3) C series exclusion was disabled as TMTPro reagent was not enabled in C-series exclusion node. The cycle time was set at 3 s and the dynamic exclusion time was set at 15 s.

APEX2-based RNAPII proximity labelling and affinity enrichment of biotinylated proteins

CRISPR–Cas9 technology was used to target endogenous *Rpb1* at the 5′ end with a cassette encoding a Flag affinity-tag and APEX2 (Flag–APEX2), resulting in RPB1 fused at its N terminus to Flag–APEX2. DPY30–mAID OsTIR1 E14 cells were co-transfected with espCas9 plasmid and a donor plasmid containing the puromycin-resistance selection gene, P2A self-cleavage site and Flag–APEX2 flanked by homology arms corresponding to the respective target genes. The sequences of the guide RNA and homology arms for targeting *Rpb1* are provided in Supplementary Table 1. The APEX2-expressing cells were incubated with 4 mM biotin-phenol reagent (Iris Biotech, LS-3500.1000) for 2 h before the start of the labelling reaction. The cells were washed with PBS (with Ca^{++} and Mg^{++}) and the labelling reaction was initiated by adding 1 mM H_2O_2 in PBS for 2 min at room temperature. The reaction was terminated by washing the cells three times with a quencher solution containing 10 mM sodium azide, 10 mM sodium ascorbate and 5 mM Trolox in PBS.

Cell fractionation for SILAC chromatin MS

Chromatin fractions were prepared as described previously[53] with some modifications. In brief, cells were lysed by swelling and mechanical force in buffer A (10 mM ammonium bicarbonate pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, 10 mM sodium ascorbate, 5 mM Trolox, 10 mM sodium azide, 1× protease inhibitor cocktail and 0.2% NP40). Nuclei were then collected by centrifugation and chemically lysed in buffer C (20 mM ammonium bicarbonate pH 8.0, 420 mM $NaCl_2$, 20% (v/v) glycerol, 2 mM $MgCl_2$, 0.2 mM EDTA, 0.1% NP40, 10 mM sodium ascorbate, 5 mM Trolox, 10 mM sodium azide, 1× protease inhibitor cocktail and 0.5 mM DTT). Lysates

were centrifuged at 20,800*g* for 45 min at 4 °C. The pellet contains the insoluble chromatin fraction and consists of DNA and proteins tightly bound to chromatin. To solubilize the chromatin pellet, 750 U Benzonase (Sigma-Aldrich) was added, followed by 10 min incubation on ice and 5 min of agitation at room temperature. Clarified lysate was collected and the protein concentration was quantified using Bio-Rad Bradford's reagent. Approximately 4 mg lysates from SILAC heavy or light cells were mixed 1:1 and incubated with 50 µl Streptavidin magnetic beads (Pierce, 88817) at 4 °C on a rotating wheel overnight. The beads were washed four times with RIPA buffer.

Sample preparation for SILAC/MS

Eluates from biotin pull-down were transferred to fresh microfuge tubes. NuPAGE sample loading buffer was added to the beads and heated at 90 °C for 5 min. A magnetic rack was used to separate the beads from the proteins. The supernatant was then run on an SDS–PAGE gel (Bis-Tris, 4–12%) enough to get the sample into the gel. Gel sections were excised, washed, reduced with DTT, alkylated with iodoacetamide and digested overnight with trypsin at 37 °C (ref. [54]). Homemade C18 StageTips were prepared as described previously[55] and preconditioned with a 50 µl wash of methanol, 50 µl wash of 70% acetonitrile/0.1% trifluoroacetic acid and two 50 µl washes of 0.1% trifluoroacetic acid at 1,000*g*. Peptides were then loaded onto StageTips and washed with 50 µl of 0.1% formic acid and were eluted with 60 µl of 70% acetonitrile/0.1% formic acid. The samples were then vacuum centrifuged using the SpeedVac and reconstituted in 0.1% formic acid for LC–MS/MS and were analysed by microcapillary LC–MS/MS using the nanoAcquity system (Waters) with a 100 µm inner-diameter × 10 cm length C18 column (1.7 µm BEH130, Waters) configured with a 180 µm × 2 cm trap column coupled to a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). Peptides were eluted at 300 nl min^{-1} using a 4 h acetonitrile gradient (0.1% formic acid). The Q-Exactive Plus mass spectrometer was operated in automatic, data-dependent MS/MS acquisition mode with one MS full scan (380–1,600 m/z) at 70,000 mass resolution and up to ten concurrent MS/MS scans for the ten most intense peaks selected from each survey scan. Survey scans were acquired in profile mode and MS/MS scans were acquired in centroid mode at 17,500 resolutions with an

isolation window of 1.5 amu and normalized collision energy of 27; AGC was set to 1×10^6 for MS1 and 5×10^4 and 50 ms max IT for MS2; charge exclusion of unassigned, +1 and greater than 6 was enabled with dynamic exclusion of 15 s.

TT_{chem}-seq and DRB/TT_{chem}-seq analysis

We performed TT_{chem}-seq as previously described[32]. In brief, cells in a 10 cm dish at 80% confluency were treated in biological duplicates at the specified time points. After the specified treatment, we supplemented the treatment medium with 1 mM 4SU and metabolically labelled the cells for 10 min. The cells were lysed in QIAzol (Qiagen, 79306) and total RNA was isolated according to the manufacturer's instructions before the addition of 100 ng of RNA spike-in mix together with QIAzol. The RNA spike-in was extracted from *Drosophila* S2 cells using 4SU, metabolically labelling the cells for 20 min. The 100 µg RNA (in a total volume of 100 µl) was fragmented by addition of 20 µl of 1 M NaOH and left on ice for 20 min. Fragmentation was stopped by addition of 160 µl of 0.5 M Tris pH 6.8 and cleaned up twice using the Micro Bio-Spin P-30 Gel Columns (BioRad, 7326223) according to the manufacturer's instructions. Biotinylation of 4SU-residues was performed in a total volume of 250 µl, containing 10 mM Tris-HCl pH 7.4, 1 mM EDTA and 5 mg of MTSEA biotin-XX linker (Biotium, BT90066) for 30 min at room temperature in the dark. RNA was then purified by phenol–chloroform extraction, denatured by 10 min incubation at 65 °C and added to 200 µl µMACS Streptavidin MicroBeads (Milentyl, 130-074-101). RNA was incubated with beads for 15 min at room temperature and beads were applied to a µColumn in the magnetic field of a µMACS magnetic separator. The beads were washed twice with pull-out wash buffer (100 mM Tris-HCl, pH 7.4, 10 mM EDTA, 1 M NaCl and 0.1% Tween-20). Biotinylated RNA was eluted twice by addition of 100 mM DTT and cleaned up using the RNeasy MinElute kit (Qiagen, 74204) using 1,050 µl ethanol (≥99%) per 200 µl reaction after addition of 700 µl of RLT buffer to precipitate RNA of less than 200 nucleotides. A total of 200 ng of the purified 4SU-labelled RNA was then used as input for the TruSeq Stranded Total RNA kit (Illumina, 20020596) for library preparation. The libraries were amplified according to the manufacturer's instructions with modifications as previously described[32]. The library was amplified with 10

PCR cycles and quality-control checked on the TapeStation (Agilent) using the High Sensitivity DNA kit before pooling and paired-end sequencing on the NextSeq 550 (Illumina) system.

For DRB/TT_{chem}-seq, cells were incubated in 100 µM DRB (Sigma-Aldrich, D1916) for 3.5 h. The cells were then washed twice in PBS, and the prewarmed fresh DRB-free medium was added to restart transcription. The RNA was labelled in vivo with 1 mM 4SU for 10 min before the addition of QIAzol, which was used to stop the reaction at the desired time point.

mNET–seq

We performed mNET–seq with minor modifications to the original protocol[27]. The Flag epitope tag was added to the N terminus of the first RNAPII subunit (RPB1) in DPY30–mAID degron cells (RPB1–APEX2 cells). Cells were seeded the day before the experiment to get 100 million cells per sample the next day. We randomly assigned flasks for each treatment and treated cells with DMSO only or auxin ligand, before extracting the chromatin-bound RNAPII. The cells were first washed with ice-cold DPBS, resuspended in 4 ml of ice-cold HLB + N buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 2.5 mM $MgCl_2$, 0.5% (v/v) NP-40 and 1× proteinase inhibitor) and incubated on ice for 5 min, then cells were scraped to a 15 ml centrifugate tube. The cell suspension was then underlaid with 1 ml of HLB + NS buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 2.5 mM $MgCl_2$, 0.5% (v/v) NP-40, 10% (w/v) sucrose and 1× proteinase inhibitor) and centrifuged to pellet the nuclei at 400*g* at 4 °C. The supernatant and membrane debris were then removed, and the nuclei were resuspended in 125 µl of NUN1 lysis buffer (20 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM EDTA, 50% (v/v) glycerol and 1× proteinase inhibitor) to which we added 1.2 ml NUN2 buffer (20 mM HEPES-KOH pH 7.6, 300 mM NaCl, 0.2 mM EDTA, 7.5 mM $MgCl_2$, 1% (v/v) NP-40, 1 M urea and 1× proteinase inhibitor) to precipitate the chromatin, and the sample was incubated on ice for 15 min with occasional vortexing. The lysates were then centrifuged at 16,000*g* for 10 min at 4 °C to pellet the chromatin. The chromatin pellets were then washed with 1× MNase buffer and then digested with 50 U MNase (NEB, M0247S) for 2 min at 37 °C with 1,400 rpm on a thermomixer. The digestion was stopped by adding 5 µl of 500 mM EGTA

(to a final concentration of 25 mM; Thermo Fisher Scientific, 50-255-956) and transferred onto ice. The reactions were then centrifuged at 16,000*g* for 5 min at 4 °C and the supernatant was subsequently diluted with 1 ml of NET-2 buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.05% (v/v) NP-40) per fraction and pooled per sample for the N-terminal Flag-RNAPII IP. We added 50 µl of anti-Flag M2 Affinity gel (Sigma-Aldrich, A2220) and incubated in the cold room for 1 h. This was followed by eight washes with NET-2 buffer and one wash with 500 µl of PNKT buffer containing 1× T4 polynucleotide kinase (PNK) buffer (NEB, M0201L) and 0.1% (v/v) Tween-20 (Thermo Fisher Scientific, BP337-100). The beads were incubated in 100 µl of PNK reaction mix containing 1× PNK buffer, 0.1% (v/v) Tween-20, 1 mM ATP and T4 PNK 3′ phosphatase minus (NEB, M0236L) at 37 °C for 10 min. We eluted the RNA by adding 350 µl of buffer RLT plus and 1 ng of fragmented *Drosophila* S2 mRNA spike-in to the extraction buffer. Next, short immunoprecipitated RNA fragments were size-selected (under 200 nucleotides) and purified from the eluates according to the manufacturer's protocol after the purification of miRNA from animal cells using the RNeasy Plus Mini Kit and the RNeasy MinElute Cleanup Kit (Qiagen, 74204), and eluted in 14 µl of nuclease-free water. We checked 1 µl of the eluted RNA on the TapeStation for the RNA quality and proceeded to NGS library preparation using the NEBNext Multiplex Small RNA Library Prep kit (NEB, NC0477293). Both were performed according to the manufacturer's protocol with low input material, with 14 PCR cycles for the library amplification step. We cleaned up and concentrated the DNA using the Monarch kit (NEB) and separated the library on a 6% TBE gel and performed size-selection by excising the smear between 147 and 307 nucleotides (according to the Quick-Load pBR322 DNA-MspI Digest ladder (NEB)). These libraries were quality-checked and quantified using the TapeStation. The samples were pooled and sequenced using paired-end sequencing on the NextSeq 550 (Illumina) system.

For promoter-proximal RNAPII half-life experiments determined by mNET–seq, chromatin was isolated from cells that were pretreated with DMSO or auxin and then incubated with 1 µM triptolide for 0, 5, 10, 20 or 40 min in the presence of DMSO/auxin and analysed using mNET–seq. Chromatin was digested with micrococcal nuclease (MNase, scissor) to release RNAPII engaged RNA from insoluble chromatin, and then immunoprecipitated using anti-Flag–RNAPII antibodies.

IP–MS data analysis

Data were analysed in Proteome Discoverer v.3.1. A database search was performed with the Sequest HT search engine using the Mouse UniProt database containing only reviewed entries and canonical isoforms (retrieved on 10 October 2019). Oxidation (M) was set as a variable modification, while TMT was set as fixed modification. A maximum of two missed cleavages were permitted. The precursor and fragment mass tolerances were 10 ppm and 0.6 Da, respectively. Peptide-spectrum matches (PSMs) were validated by percolator with a 0.01 posterior error probability threshold. Only PSMs with an isolation interference of <25% and at least 5 MS2 fragments matched to the peptide sequence selected for MS3 were considered. The quantification results of PSMs were combined into protein-level quantification using the MSstatsTMT R package[56]. Only proteins with at least three peptides were reported. To identify interactors, we performed differential abundance analysis between the IP samples and their corresponding controls (that is, 0 h IP was compared to 0 h IgG control and 8 h IP was compared to 8 h IgG control). A protein was considered to be an interactor if in one or both comparisons its levels were statistically significantly different ($Q \leq 0.05$, limma test, with P values adjusted by the Storey method) and at least twice higher in IP reactions than in the corresponding IG control (Supplementary Table 3).

SILAC/MS analyses

All SILAC/MS data were processed using the MaxQuant software (Max Planck Institute of Biochemistry; v.1.5.3.30). The default values were used for the first search tolerance and main search tolerance—20 ppm and 6 ppm, respectively. Labels were set to Arg10 and Lys8. MaxQuant was set up to search the reference mouse proteome database downloaded from UniProt on 9 January 2020. MaxQuant performed the search assuming trypsin digestion with up to two missed cleavages. Peptide, site and protein false-discovery rate (FDR) were all set to 1% with a minimum of one peptide needed for identification but two peptides needed to calculate a protein level ratio. The following modifications were used as variable modifications for identifications and included for protein quantification: oxidation of methionine, acetylation of the protein N terminus, ubiquitination of lysine,

phosphorylation of serine, threonine and tyrosine residues, and carbamidomethyl on cystine. Intensity values measured in all replicates were log_2-transformed (Supplementary Table 4), *P* values were computed using Fisher's tests and corrected using Benjamini–Hochberg FDR correction. All raw MS data files have been deposited to the ProteomeXchange Consortium (and PXD039176). The Gene Ontology term enrichment analysis was performed using Enrichr online tool (https://maayanlab.cloud/Enrichr/), STRING (https://string-db.org) and clusterProfiler[57].

ChIP–seq data analysis

The sequenced reads were demultiplexed using bcl2fastq (v.2.19.0.316), and basic quality control was performed on the resulting FASTQ files using FastQC (v.0.11.8). FASTQ reads were mapped to the GRCm38 (mm10) genome using Bowtie2 (v.2.4.1) using the standard settings. The resulting SAM files were converted to BAM files using the SAMtools (v.1.10) view command, after which the BAM files were sorted and indexed, and potential PCR duplicates were removed using the rmdup function. DeepTools (v.3.3.0) was used to generate occupancy heat maps, and the resulting normalized occupancy matrix was used as input for public R scripts to generate average profile plots and to calculate processivity indices. In brief, the BAM files were converted into BigWig files using the bamCoverage function (bamCoverage -p 8 --normalizeUsing RPGC --effectiveGenomeSize mm10 --centerReads -e --scaleFactor X --blackListFileName mm10.blacklist.bed). For comparison, quantitative ChIP–seq data using spike-in normalization were used, normalization to the *Drosophila* S2 spike-in was performed at this stage according to the manufacturer's instructions (Active Motif, 61686; https://www.activemotif.com/catalog/1091/chip-normalization). The computeMatrix function was used to quantify the occupancy of reads across the specified intervals, and the plotProfile and plotHeatmap functions were used to plot the data. Reproducibility of replicates is shown in Supplementary Fig. 3.

Quant-seq/SLAM-seq analysis

Gene and 3′ untranslated region (UTR) annotations were obtained from the UCSC table browser (https://genome.ucsc.edu/cgi-bin/hgTables, mm10 vM14 3′ UTR). Adapters were trimmed from raw reads using cutadapt through the trim_galore wrapper tool with adapter overlaps set to 3 bp for trimming. For Quant-seq, concatenated fastq files were trimmed for adapter sequences, and masked for low-complexity or low-quality sequences using trim_galore, then mapped to the mm10 whole genome using HISAT v.2.2.1 with the default parameters. The number of reads mapped to the 3′ UTR of genes was determined using featureCounts. Raw reads were normalized to CPM. SLAM-seq analysis was performed as previously described[58] using the SlamDunk package[59]. Trimmed reads were further processed with SlamDunk (v.0.3.4 16). The 'Slamdunk all' command was executed with the default parameters except '-rl 74 -t 8 fastq.gz -n 100 -m -mv 0.2 -o Slamdunk2', running the full analysis procedure (slamdunk all) and aligning against the mouse genome (GRCm38), filtering for variants with a variant fraction of 0.2. Unless indicated otherwise, reads were filtered for having ≥2 T>C conversions. The remaining parameters were left as defaults.

Analysis of differential gene expression was restricted to genes with ≥10 reads in at least one condition. Differential gene expression calling was performed on raw read counts with ≥2 T>C conversions using DESeq2 with the default settings, and with size factors estimated on corresponding total mRNA reads for global normalization. Downstream analysis was restricted to genes that passed all internal filters for FDR estimation by DESeq2. Plots of differential gene expression were visualized using the ggplot2 package in R with significant genes (P value < 0.05, |log2FC| ≥ 1). Reproducibility of replicates is shown in Supplementary Fig. 3.

mNET–seq data analysis

Reads were demultiplexed using bcl2fastq, then trimmed for adapter content with cutadapt (-m 10 -e 0.05 --match-read-wildcards -n 1), and mapped with STAR to the GRCm38 (mm10) genome assembly. Further data processing was performed using the R/Bioconductor environment. Coverage tracks for further analysis were restricted to the last nucleotide incorporated by the RNAPII in the aligned mNET–seq reads as described[60]. To calculate the half-life of paused RNAPII for each gene by mNET–seq, RNAPII density

was calculated in a 300 bp window downstream of the TSS. RNAPII time-course measurements were fitted into an exponential decay model using the RNAdecay R package (https://bioconductor.org/packages/release/bioc/html/RNAdecay.html). We selected genes fulfilling the current criteria: (1) detectable RNAPII levels (reads per kilobase of transcript, per million mapped reads > 1), (2) highest RNAPII density under the no triptolide (0 min) condition and (3) low variance between replicates ($\sigma < 0.05$). Genes fitting the above criteria (n = 6,338) were used to calculate the RNAPII half-life. Reproducibility of replicates is shown in Supplementary Fig. 3.

TT_{chem}-seq data analysis

Paired-end reads were demultiplexed using bcl2fastq. TT_{chem}-seq raw data were processed essentially as described previously[32]. Raw reads were aligned to the mouse mm10 genome assembly using STAR. Mapped reads with a mapping quality score <10 were discarded with SAMtools. All further processing was performed using the R/Bioconductor framework. Antisense bias, sequencing depth and cross-contamination rates were calculated as described previously[32]. Reads were mapped to transcription units, which represent the union of all annotated UCSC RefSeq isoforms per gene. The number of transcribed bases per transcription unit was calculated as the sum of the coverage of evident (sequenced) fragment parts (read pairs only) for all fragments in addition to the sum of the coverage of the inner mate interval if not entirely overlapping a RefSeq annotated intron (UCSC RefSeq GRCm38). Computational analysis DRB/TT_{chem}-seq data were processed using a previously published protocol[32]. In brief, reads were aligned to human GRCm38 (mm10). Read depth coverage was normalized to account for differences between samples using a scale factor derived from a spike-in aligned and counted against *Drosophila melanogaster*. Biological replicate alignments were combined for the purpose of visualization and wave-peak analysis to increase read-depth coverage.

A set of non-overlapping protein-coding genes of >60 kb and <300 kb was selected for wave-peak analysis. A meta-gene profile was calculated by taking a trimmed mean of each base-pair coverage in the region around the TSS. This was further smoothened using a spline. Wave peaks were called at

the maximum points on the spline, with the stipulation that the peak must advance with time before being subjected to manual review. Elongation rates (kb per min) were calculated by fitting a linear model to the wave-peak positions as a function of time. For elongation-rate analysis, the following criteria were used to filter genes: The 0 min timepoint DMSO control sample was required to show expression of the gene (mean expression of >100 rpm by TT-seq) and was required to have a wave peak called within 10 kb of the pausing peak region to remove artifacts. Genes showing an increase in transcription in the DMSO control sample for the time course were identified by requiring the wave peak in the 0 min sample to be less than the wave peak in the 10 min timepoint wave peak, and the wave peak in the 10 min sample to be less than the wave peak in the 20 min timepoint, and the wave peak in the 20 min sample to be less than the wave peak in the 30 min timepoint. This resulted in the identification of 855 genes, for which elongation rates were calculated for the samples by dividing the wave peak position by the timepoint. Reproducibility of replicates is shown in Supplementary Fig. 3.

Statistics and reproducibility

The statistical details of the experiments can be found in the figure legends and in the Methods. Western blotting in Figs. 1b,c,f–h, 3a and 5b,c,j was independently performed three times with similar results, and western blotting in Extended Data Figs. 8c,f,g and 9c,e was performed twice as biologically independent experiments.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All related raw sequencing and processed data have been deposited at the Gene Expression Omnibus under accession number GSE181714.

References

1. Lauberth, S. M. et al. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).

2. Shilatifard, A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* **81**, 65–95 (2012).

3. Miller, T. et al. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl Acad. Sci. USA* **98**, 12902–12907 (2001).

4. Qu, Q. et al. Structure and conformational dynamics of a COMPASS histone H3K4 methyltransferase complex. *Cell* **174**, 1117–1126 (2018).

5. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780 (2016).

6. Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science* **357**, eaal2380 (2017).

7. Rao, R. C. & Dou, Y. Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* **15**, 334–346 (2015).

8. Vermeulen, M. et al. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**, 967–980 (2010).

9. Ooi, S. K. et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717 (2007).

10. Liu, X. Y. et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**, 558–562 (2016).

11. Sankar, A. et al. KDM4A regulates the maternal-to-zygotic transition by protecting broad H3K4me3 domains from H3K9me3 invasion in oocytes. *Nat. Cell Biol.* **22**, 380–388 (2020).

12. Chen, K. F. et al. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat. Genet.* **47**, 1149–1157 (2015).

13. Clouaire, T. et al. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev.* **26**, 1714–1728 (2012).

14. Ang, Y. S. et al. Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183–197 (2011).

15. Jiang, H. et al. Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* **144**, 513–525 (2011).

16. Denissov, S. et al. Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* **141**, 526–537 (2014).

17. Dorighi, K. M. et al. Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. *Mol. Cell* **66**, 568–576 (2017).

18. Bledau, A. S. et al. The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Development* **141**, 1022–1035 (2014).

19. Howe, F. S., Fischl, H., Murray, S. C. & Mellor, J. Is H3K4me3 instructive for transcription activation. *Bioessays* **39**, 1–12 (2017).

20. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep.* **15**, 210–218 (2016).

21. Nabet, B. et al. The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* **14**, 431–441 (2018).

22. Christensen, J. et al. RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. *Cell* **128**, 1063–1076 (2007).

23. Schmitz, S. U. et al. Jarid1b targets genes regulating development and is involved in neural differentiation. *EMBO J.* **30**, 4586–4600 (2011).

24. Herzog, V. A. et al. Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Methods* **14**, 1198–1204 (2017).

25. Vermeulen, M. et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).

26. Rahl, P. B. et al. c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).

27. Nojima, T., Gomes, T., Carmo-Fonseca, M. & Proudfoot, N. J. Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat. Protoc.* **11**, 413–428 (2016).

28. Shao, W. & Zeitlinger, J. Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.* **49**, 1045–1051 (2017).

29. Mylonas, C., Lee, C., Auld, A. L., Cisse, I. I. & Boyer, L. A. A dual role for H2A.Z.1 in modulating the dynamics of RNA polymerase II initiation and elongation. *Nat. Struct. Mol. Biol.* **28**, 435–442 (2021).

30. Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* **3**, e02407 (2014).

31. Gressel, S. et al. CDK9-dependent RNA polymerase II pausing controls transcription initiation. *eLife* **6**, e29736 (2017).

32. Gregersen, L. H., Mitter, R. & Svejstrup, J. Q. Using TT_{chem}-seq for profiling nascent transcription and measuring transcript elongation. *Nat. Protoc.* **15**, 604–627 (2020).

33. Caizzi, L. et al. Efficient RNA polymerase II pause release requires U2 snRNP function. *Mol. Cell* **81**, 1920–1934 (2021).

34. Zumer, K. et al. Two distinct mechanisms of RNA polymerase II elongation stimulation in vivo. *Mol. Cell.* **81**, 3096–3109 (2021).

35. Geijer, M. E. et al. Elongation factor ELOF1 drives transcription-coupled repair and prevents genome instability. *Nat. Cell Biol.* **23**, 608–619 (2021).

36. Levine, M. Paused RNA polymerase II as a developmental checkpoint. *Cell* **145**, 502–511 (2011).

37. Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456–475 (2016).

38. Ji, X. et al. Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proc. Natl Acad. Sci. USA* **112**, 3841–3846 (2015).

39. Chen, F. X. et al. PAF1, a molecular regulator of promoter-proximal pausing by RNA polymerase II. *Cell* **162**, 1003–1015 (2015).

40. Yu, M. et al. RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science* **350**, 1383–1386 (2015).

41. Zheng, H. et al. Identification of integrator-PP2A complex (INTAC), an RNA polymerase II phosphatase. *Science* **370**, eabb5872 (2020).

42. Gardini, A. et al. Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell* **56**, 128–139 (2014).

43. Tatomer, D. C. et al. The integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev.* **33**, 1525–1538 (2019).

44. Elrod, N. D. et al. The integrator complex attenuates promoter-proximal transcription at protein-coding genes. *Mol. Cell* **76**, 738–752 (2019).

45. Beckedorff, F. et al. The human integrator complex facilitates transcriptional elongation by endonucleolytic cleavage of nascent transcripts. *Cell Rep.* **32**, 107917 (2020).

46. Dasilva, L. F. et al. Integrator enforces the fidelity of transcriptional termination at protein-coding genes. *Sci. Adv.* **7**, eabe3393 (2021).

47. Vervoort, S. J. et al. The PP2A–integrator–CDK9 axis fine-tunes transcription and can be targeted therapeutically in cancer. *Cell* **184**, 3143–3162 (2021).

48. Barbieri, E. et al. Targeted enhancer activation by a subunit of the integrator complex. *Mol. Cell* **71**, 103–116 (2018).

49. Copenhaver, G. P. et al. ATX1-generated H3K4me3 is required for efficient elongation of transcription, not initiation, at ATX1-regulated genes. *PLoS Genet.* **8**, e1003111 (2012).

50. Hughes, A. L. et al. A CpG island-encoded mechanism protects genes from premature transcription termination. *Nat. Commun.* **14**, 726 (2023).

51. Zheng, H. et al. Structural basis of INTAC-regulated transcription. Preprint at *bioRxiv* https://doi.org/10.1101/2021.11.29.470345 (2021).

52. Stein, C. B. et al. Integrator endonuclease drives promoter-proximal termination at all RNA polymerase II-transcribed loci. *Mol. Cell* **82**, 4232–4245 (2022).

53. Spruijt, C. G., Baymaz, H. I. & Vermeulen, M. Identifying specific protein–DNA interactions using SILAC-based quantitative proteomics. *Gene Reg.* **977**, 137–157 (2013).

54. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2006).

55. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).

56. Choi, M. et al. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).

57. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).

58. Muhar, M. et al. SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science* **360**, 800–805 (2018).

59. Neumann, T. et al. Quantification of experimentally induced nucleotide conversions in high-throughput sequencing datasets. *BMC Bioinform.* **20**, 258 (2019).

60. Prudencio, P., Rebelo, K., Grosso, A. R., Martinho, R. G. & Carmo-Fonseca, M. Analysis of mammalian native elongating transcript sequencing (mNET-seq) high-throughput data. *Methods* **178**, 89–95 (2020).

Acknowledgements

We thank H. Damhofer, K. Nishimura, D. Shlyueva, S. Sidoli, Z. Li, Z. Sun, Y. Lin and Y. Fang for technical advice and reagents; C. Huang, A. Radzisheuskaya, D. Shlyueva and R. Armstrong for comments on the manuscript; and members of the Helin laboratory for discussions and support. This work was supported by the Memorial Sloan Kettering Cancer Center Support Grant (no. NIH P30 CA008748), a Tri-Institutional Stem Cell grant (no. 2019-035), startup funds from The Institute of Cancer Research and the Novo Nordisk Foundation to the NNF Center for Stem Cell Biology (no. NNF17CC0027852).

Author information

Authors and Affiliations

1. Cell Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

   Hua Wang, Xuejun Jiang & Kristian Helin

2. Center for Epigenetics Research, Memorial Sloan Kettering Cancer Center, New York, NY, USA

   Hua Wang, Xuejun Jiang & Kristian Helin

3. The Institute of Cancer Research, London, United Kingdom

   Zheng Fan & Kristian Helin

4. Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark

   Zheng Fan & Kristian Helin

5. The Novo Nordisk Foundation Center for Stem Cell Biology (Danstem), University of Copenhagen, Copenhagen, Denmark

   Zheng Fan & Kristian Helin

6. Microchemistry and Proteomics Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA

   Pavel V. Shliaha, Matthew Miele & Ronald C. Hendrickson

Contributions

H.W. and K.H. conceived the project. H.W. designed and performed the experiments and bioinformatics analysis. P.V.S. and M.M. performed the MS studies, supervised by R.C.H. Z.F. and X.J. provided input to the project.

H.W. and K.H. wrote the manuscript. K.H. supervised the study and acquired funding.

Corresponding author

Correspondence to Kristian Helin.

Ethics declarations

Competing interests

K.H. is a co-founder of Dania Therapeutics, consultant for Inthera Bioscience and a scientific advisor for MetaboMed and Hannibal Innovation. The other authors declare no competing interests.

Peer review

Peer review information

Nature thanks Alvaro Rada-Iglesias and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Generation of DPY30–mAID and RBBP5–FKBP mES cell lines.

(**a**,**b**) DPY30, RBBP5, H3K4me3 and RNAPII occupancy in mES cells plotted around the TSS of all protein-coding genes. Genes were sorted based

on H3K4me3 binding levels in the heat maps (**a**). 3′ RNA-seq data is from Quant-seq in mES cells. The average signal in the profiles (**b**) is plotted over the transcription start sites (TSS ± 2 kb) of all protein-coding genes. (**c**) Outline of CRISPR-HDR based knock-in targeting approach generating DPY30-AID and RBBP5–FKBP, by knocking-in mAID and FKBP12 degron tags into the 3′ end of both alleles of endogenous *Dpy30* and *Rbbp5* loci of the E14 mES cell cell line, respectively. The endogenous *Dpy30* gene was edited to encode Auxin inducible degradation (mAID)-BFP at the C terminus of DPY30 in OsTir1 E14 mES cells. The endogenous *Rbbp5* gene was edited to encode FKBP12 tag-Neomycin that enables targeted degradation upon dTAG-13 treatment. (**d**) DPY30 expression in the indicated cell lines as measured by flow cytometry analysis using an anti-DPY30 antibody. (**e,f**) H3K4me3 ChIP–seq heat maps and profiles at H3K4me3 peak centre in control and Auxin-treated cells in DPY30–mAID (E) or dTAG-13-treated in RBBP5–FKBP (F) cells. The signal was plotted over H3K4me3 peaks centre (peak centre ± 5 kb) which are called from WT mES cells. Sites are sorted by the ChIP–seq signals at 0 h. (**g,h**) Integrative genomics viewer (IGV) browser snapshots comparing DPY30 (g) or H3K4me3 (h) enrichments determined by ChIP–seq in control and Auxin-treated in DPY30–mAID degron mES cells. (**i–k**) ChIP-qPCR enrichments of DPY30 (I) and H3K4 methylations (j and k) in DPY30–mAID cells after treatment with Auxin for the indicated times. (**l**) ChIP-qPCR for H3K4me3 in RBBP5–FKBP cells. In panels (i-l) target sites around the promoter of the genes and control region (intragenic chr8: 72,806,101- 72,806,240) were used. For H3K4me1, the enhancer region of *Nanog* was used for DPY30, H3K4me2 and H3K4me3 ChIP-qPCR. Graph shows mean values from technical triplicates (n = 3), from one representative out of two independent experiments.

Extended Data Fig. 2 DPY30 and RBBP5 are required for cell proliferation.

(**a**) Cell proliferation assays of DPY30–mAID and RBBP5–FKBP cells grown either with or without Auxin/dTAG-13. The line graph represents the mean ± SD of the numbers of cells counted at each time point. Parental represents the parental E14 ES cells without CRISPR editing. Ectopic expression (EE) of DPY30 or RBBP5 rescues the proliferation of defect

observed by degrading endogenous DPY30 and RBBP5, respectively. Data are from three biological replicates (n = 3) and were analysed using Two-way ANOVA (n = 3) and represented as mean ± s.d., $***P < 0.001$. (**b**) Colony formation assay for DPY30–mAID and RBBP5–FKBP cells grown either with or without Auxin/dTAG-13 for two weeks. (**c**) Quantification of colony formation assays from two independent experiments. Data in bar plots are represented as mean and n = 2 replicates. (**d**) Cell proliferation assays of DPY30–mAID cells after treatment with Auxin at the indicated times with or without *Kdm5ab* dKO. (**e**) RT-qPCR analyses of mRNA expression of DPY30–mAID cells with or without *Kdm5ab* dKO. Two independent dKO clones were chosen for downstream analysis. The values are normalized to 18S rRNA. Graph shows mean values from technical triplicates (n = 3), from one representative out of two independent experiments. (**f**) ChIP-qPCR enrichment for WDR5, SETD1B and MLL1 in DPY30–mAID cells system after treatment with Auxin at the indicated points. Graph shows mean values from technical triplicates (n = 3), from one representative out of two independent experiments.

Extended Data Fig. 3 Outline and controls for SLAM-seq experiments.

(**a**) Experimental design. To validate the SLAM-seq protocol, we performed a pilot experiment for mapping responses to short-term THZ1 (2h) treatment by SLAM-seq in mES cells. SLAM-seq utilizes thymine-to-cytosine (T>C) conversion from 4-thiouridine (4sU)-labelled mRNAs to quantify the abundance of nascent RNA transcripts using 3′ end mRNA-sequencing (Quant-seq). To monitor the consistency and reproducibility of different SLAM-seq data, we inhibited transcription with THZ1 (reduces RNAPII-mediated gene transcription by inhibiting cyclin-dependent kinase 7 (CDK7)). (**b**) Conversion rates for each position of 4-thioU-containing SLAM-seq reads (≥ 2 T>C conversions) before or after Auxin or THZ1 treatment for 2 h. Changes in the abundance of newly synthesized mRNAs (detected in SLAM-seq based on T>C conversions). Average conversion rates (centre line) ± s.d. (whiskers) of two independent experiments (points) are shown. *P* value (Two-sided Mann-Whitney test) is indicated ($***P <$ 0.001, n.s., not significant.). n = 20,428 transcripts. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers

show the 1.5× IQR above and below the box. (**c,d**) Transcriptional response of the cells treated with THZ1/DMSO for 2h followed by 4sU labelling over 60 min. (c) MA plots comparing total gene expression level with log change in transcription per gene measured by 4-thioU RNA-seq (SLAM-seq). (d) MA plots comparing nascent gene expression levels with log change in transcription per gene measured by SLAM-seq. THZ1 treatment confirmed that transcripts containing T>C conversions of protein-coding genes were broadly repressed, which captured the prominent immediate responses, while the total mRNA level showed fewer changes. *P*-adjusted value by Wald test in DESeq2. (**e**) H3K4me3 levels on TSS before Auxin treatment of downregulated genes (n = 1,111) and unchanged genes (n = 10,107) measured by SLAM-seq (in response to Auxin treatment for 2h in DPY30–mAID cells). The *p* value was calculated with a two-sided Wilcoxon test. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers show the 1.5× IQR above and below the box. (**f**) H3K4me3 peak width of downregulated genes and unchanged genes measured by ChIP–seq at steady-state. n= for downregulated and n= for unchanged genes. The *p* value was calculated with a two-sided Wilcoxon test. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers show the 1.5× IQR above and below the box. (**g**) Box plot showing log2-transformed fold change of nascent transcription (SLAM-seq, 2 h vs. 0 h) of genes containing CGI (n = 11,386) and non-CGI (n = 3,262) promoters. The *p* value was calculated with a two-sided Wilcoxon test. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers show the 1.5× IQR above and below the box. (**h**) Gene set enrichment analysis (GSEA) of downregulated genes in DPY30–mAID cells in response to 2 h Auxin treatment.

Extended Data Fig. 4 H3K4me3 loss does not have detectable effects on binding of TAF3, CDK7 and TBP to transcription start sites and to PIC formation.

(**a**) ChIP–seq profiles of the indicated proteins using DPY30–mAID and RBBP5–FKBP cells treated with or without Auxin/dTAG-13, respectively for 8 h. The enrichments were plotted over the transcription start sites (TSS ± 5 kb) of protein-coding genes. TSS, transcription start site. (**b**) Outline of the mass spectroscopy proteome profiling strategy for mapping the

interaction networks of RNAPII in DPY30–mAID cells treated with or without Auxin. (**c**) String network of protein complexes (*k*-means clustering) showing RNAPII interactors in control cells compared with IgG mock IP. (**d**) Volcano plot showing proteins changing their association with RNAPII in response to acute loss of H3K4me3. The *x* axis displays the enrichment (log2 fold change) of proteins in Auxin-treated cells (Auxin 8 h) compared to DMSO-treated control cells (Auxin 0 h). The *y* axis shows the significance (-log10 *P* value) of enrichment calculated from three biological replicate experiments. A protein was considered an interactor if in one or both comparisons its levels were statistically significantly different (Q value ≤ 0.05, limma test, with *P* values adjusted by the Storey method). (**e**) RNAPII occupancy based on ChIP–seq in the indicated cell lines. Metagene analysis showing the genome-wide enrichment averages on protein-coding genes, data are shown along with 3 kb upstream of the transcriptional start site to 5 kb downstream of the end of each annotated gene. TSS, transcription start site, TES, transcription end site. (**f**) Estimation of a gene's "pausing index" (PI) from RNAPII ChIP–seq data. The promoter is defined as the region covering 200 bp upstream to 200 bp downstream of the TSS; the gene body is defined as the region from 400 bp downstream of the TSS to TES, genes with gene length less than 400 bp are removed for pausing index analysis. (**g**) Violin plots showing changes of gene expression in the indicated samples. Genes were separated into three equal parts based on their accumulation change of RNAPII in promoter-proximal region. (**h**) An IGV snapshot comparing RNAPII, RNAPII Ser 2p and Ser 5p ChIP–seq signals in control and Auxin-treated DPY30–mAID cells at the indicated times. (**i**) Average metagene ChIP–seq profiles for the indicated factors in control and dTAG-13-treated RBBP5–FKBP cells.

Extended Data Fig. 5 The fast turnover of H3K4me3 is dependent on KDM5 demethylases.

(**a**) RNAPII pausing index in DPY30–mAID (black), DPY30–mAID: *Kdm5a/b* dKO (blue, clone #2) and Auxin-treated cells. Higher index values indicate a higher degree of RNAPII pausing. (**b**) ChIP-qPCR enrichment of H3K4me3 in DPY30–mAID cells after treatment with Auxin at the indicated times with or without *Kdm5a/b* dKO. (**c**) ChIP-qPCR signals for RNAPII in DPY30–mAID cells following treatment with Auxin at the indicated times

with or without *Kdm5a/b* dKO. For ChIP-qPCR, the target sites around promoter of the indicated genes and control region (intragenic chr8: 72,806,101- 72,806,240) were used. Graph shows mean values from technical triplicates (n = 3), from one representative out of two independent experiments.

Extended Data Fig. 6 H3K4me3 regulates the paused RNAPII half-life.

(**a**) Half-lives of paused RNAPII in control and Auxin-treated DPY30–mAID degron cells. The half-life was calculated based on an exponential decay model. Normalized promoter-proximal RNAPII density for each gene is shown over the course of triptolide treatment both for control (DMSO, black) and Auxin (blue) conditions. n = 2 biological replicates. (**b**) Boxplot showing increased paused RNAPII duration following H3K4me3 loss. *P* value was calculated with a two-sided Wilcoxon test, n = 4,007 genes. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers show the 1.5× IQR above and below the box. (**c**) An IGV snapshot comparing mNET–seq signals after triptolide-induced block of transcription in control and Auxin-treated in DPY30–mAID cells at the indicated time points. The decreasing levels of paused RNAPII were observed in the pausing window (shaded areas) over time. (**d**) Heat maps and profiles showing changes in mNET–seq signals upon acute loss of H3K4me3. Heat maps were plotted as increasing gene length over regions 5 kb upstream to 50 kb downstream of the TSS. The right panel display the difference (Δ) in mNET–seq between control (+ Auxin 0 h) and Auxin-treated (+ Auxin 8 h) cells. Genes were sorted by gene length. (**e**–**f**) Heat maps and profiles showing changes of H3K4me3 ChIP–seq and mNET–seq at promoters, active and inactive enhancers upon Auxin treatment (8 h) in DPY30–mAID degron cells. The signals were plotted over the transcription start sites (TSS ± 5 kb) or the centre of enhancers (centre ± 5 kb). (**g**) H3K36me3 ChIP–seq profiles in control (+ Auxin 0 h) and Auxin-treated (+ Auxin 8 h) cells in DPY30–mAID cells. Genes were split by their H3K4me3 and H3K27me3 levels around TSS regions. (**h**) Histogram of RNAPII elongation rates for individual genes between 60 and 300 kb with RPM value ≥100 across all time points (n = 855) with a 10-min wave peak called beyond 2 kb and sequential increase from the TSS over the 10-, 20- and 30-

min time points. (**i**) Heat maps and profiles showing changes of ChIP–seq, mNET–seq and TT_{chem}-seq at promoters upon Auxin treatment (8 h) in DPY30–mAID degron cells. The downregulated genes and unchanged genes measured by SLAM-seq. The signals were plotted over the transcription start sites (TSS ± 5 kb). (**j**) Boxplot showing RNAPII elongation rates following H3K4me3 loss. *P* value was calculated with a two-sided Wilcoxon test. The number of genes (n) from each group are shown at the bottom. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers show the 1.5× IQR above and below the box.

Extended Data Fig. 7 Loss of H3K4me3 does not have detectable effects on transcriptional initiation.

(**a**) Experimental strategy for retinoic acid (RA) differentiation in control and Auxin-treated degron cells. (**b**) Principal component (PC) analysis of RA-induced mRNA expression changes in DPY30–mAID cells prior to treatment with Auxin (circles) or DMSO (triangles) for 2 h at the indicated differentiation points: 0h, 8h, d1, d2, d4, d6. Developmental trajectory is shown by the dashed arrow. (**c**) Spearman correlation heat map of retinoic acid (RA) time course RNA-seq replicates of Auxin-treated and DMSO-treated cells. (**d**) Short- and long-term effects of DPY30 degradation on the expression of RA-induced genes. DPY30–mAID cells were treated with DMSO, Auxin and RA as indicated. The stable RA induced genes were identified from upregulated genes at day 6 in control cells (DMSO). The early RA induced genes were identified from upregulated genes at 8 h in control cells (DMSO). Data were analysed using Two-way ANOVA (n = 2) and represented as mean ± s.d., $^{**}P < 0.01$, $^{***}P < 0.001$. (**e**) Gene set enrichment analysis (GSEA) analysis of RA induced genes (RA upregulated gene list from previously published data, n = 227). The curve represents the evolution of the density of the genes identified in the RNA-seq. The False Discovery Rate (FDR) is calculated by comparing the actual data with 1000 Monte-Carlo simulations. The NES (Normalized Enrichment Score) computes the density of modified genes in the dataset with the random expectancies, normalized by the number of genes found in the given gene cluster, considering the size of the cluster. (**f,g**) H3K4me3 and RNAPII occupancy before or after RA treatment in control and Auxin-treated degron cells. The enrichments were plotted over the transcription start sites (TSS ±

2 kb) of protein-coding genes. Rows are sorted by decreasing H3K4me3 ChIP–seq occupancy in the control (DMSO RA 0h) cells.

Extended Data Fig. 8 APEX2-based proteomic mapping scheme and characterization of endogenous RNAPII *in vivo* interactions.

(**a**) Genomic confirmation of the APEX-modified *Rpb1* loci in DPY30–mAID and RBBP5–FKBP degron cells. (**b**) Sanger sequencing of the wild type and Flag-APEX2-RPB1 knock-ins. (**c**) Western blot showing the expression of APEX-modified RPB1 in DPY30–mAID and RBBP5–FKBP degron cells. (**d**) Representative brightfield images of mES cell colonies. (**e**) RT-qPCR analysis of the expression of pluripotency and differentiation genes in the knock-in cells. Data are from three biological replicates ($n = 3$) and are analysed using Two-way ANOVA and represented as mean ± s.d. (**f**) Titration of biotin phenol (BP). Cells stably expressing Flag-APEX2-RPB1 were pre-incubated for 30 min with the indicated concentrations of BP, followed by the addition of 1 mM H2O2 for 1 min. Cell lysates were probed with Streptavidin-HRP. Proximity biotinylation is optimal at a BP concentration of 4 mM. (**g**) Confirmation of the APEX2 functionality by protein biotinylation in the APEX2-engineered DPY30–mAID and RBBP5–FKBP degron cells. The discrete bands, denoted with asterisks, show APEX2-independent biotinylation by native enzymes. The Connexin-APEX2 overexpressed cell was severed as positive control for the APEX2 system. (**h**) SILAC-based chromatin proteomic strategy for mapping the neighbourhood interaction networks of APEX2-tagged RNAPII. (**i**) Principal component analysis (PCA) of SILAC signal in the RNAPII-APEX2 cells with or without agonist. (**j**) Distribution of SILAC ratio of RNAPII interactions quantified in the chromatin proteomic analyses. Mean log2 SILAC ratio is shown. In total, 1,901 proteins were identified in this experiment. The RNAPII-APEX2-bait (BP+H2O2) population has a right-shifted distribution compared with the no agonist negative control population, which indicates that the log2(SILAC) ratio allows us to distinguish bona fide RNAPII interactions from non-RNAPII interactions. (**k**) GO network showing significantly (q value < 0.001) enriched terms for positive RNAPII interaction neighbourhoods from RNAPII-APEX2 experiment. The most prominent pathways are indicated. Connecting lines

show interaction of protein nodes. (**l**) KEGG enrichment and GO network showing significantly (q value < 0.001) enriched terms for positive RNAPII interaction neighbourhoods from RNAPII-APEX2 experiment. (**m**) Principal component analysis (PCA) of SILAC signal in the RNAPII-APEX2 DPY30–mAID cells with or without Auxin treatment. Time trajectory is shown by the dashed arrow. (**n**) Venn diagram indicating overlap between up or downregulated targets in the indicated samples. (**o**) Heatmap representing relative protein abundance of DPY30 and selected targets in Auxin treated DPY30–mAID cells. n = 3 independently samples. (**p**) Scatterplot analysis of proteins identified by SILAC in RNAPII-APEX2 DPY30–mAID cells following Auxin treatment for 2 and 8 h. (**q**) Gene ontology-based functional classification of 228 downregulated proteins in RNAPII-APEX2 DPY30–mAID cells following Auxin treatment for 2 and 8 h. The dot size is proportional to the number of members in an enrichment set, and colour intensity reflects the p value. Significance based on clusterProfiler analysis with Benjamini-Hochberg-adjusted P values.

Extended Data Fig. 9 INTS11 is required for nascent transcription.

(**a**) Venn diagram indicating overlap of H3K4me3 interactors and RNAPII-APEX2 dependent interactors from ChIP-MS (chromatin proteomic profiling) data. (**b**) Relative enrichments of selected targets in various ChIP preparations based on ChIP-MS. (**c**) Validation of INTS11 interaction with H3K4me3 in RBBP5–FKBP degron cells at different times after dTAG-13 addition. Biotinylated proteins within lysates were enriched using Streptavidin-coated magnetic beads and analysed by Western blot. In parallel, sample in which H_2O_2 was omitted was prepared as negative control. (**d**) Schematic representation of the dTAG INTS11 targeting strategy for the INTS11–FKBP degron mES cells. (**e**) Western blot showing the expression of INTS11 and INTS11–FKBP–HA, using antibodies recognizing INTS11 or the HA in parental and knock-in degron cells. The arrow indicates the specific HA-tagged INTS11–FKBP–HA protein. (**f**) RT-qPCR analysis showing the expression of selected pluripotency and differentiation genes in the parental and INTS11–FKBP knock-in cells. Data are from three biological replicates (n = 3) and are analysed using Two-way ANOVA and represented as mean ± s.d. (**g**) Growth curve analysis of

parental and INTS11–FKBP E14 cells treated with or without dTAG-13. (**h**) INTS11 enrichment profiles and heat maps as determined by using the HA-tag in control (0 h) and Auxin-treated (2 h) DPY30–mAID; INTS11–FKBP degron cells. Genome-wide binding averages showed enrichments at the TSS regions (TSS ± 2 kb) of protein coding genes. TSS, transcription start site. Rows were sorted by decreasing ChIP–seq occupancy in the control (0 h) cells. (**i**) Correlations between TT_{chem}-seq replicate experiments in INTS11–FKBP degron cells treated with or without dTAG-13 for the indicated times. (**j**) Average profiles for TT_{chem}-seq for the upstream anti-sense RNAs of each annotated protein-coding gene in INTS11 degron cells. TSS, transcription start site. (**k**) RNAPII profiles of various subclasses of annotations in INTS11–FKBP degron cells with or without dTAG-13 treatment. TSS, transcription start site. mRNA, messenger RNA. snRNA, small nuclear RNA. ncRNA, non-coding RNA. eRNA, enhancer RNA.

Extended Data Fig. 10 Loss of INTS11 causes reduced transcriptional output of protein-coding genes.

(**a**) Experimental design of SLAM-seq for INTS11–FKBP degron cells. Conversion rates for each position of a 4-thioU-containing SLAM-seq reads (≥ 2 T>C conversions) before or after dTAG-13 treatment. Average conversion rates (centre line) ± s.d. (whiskers) are shown. n = 3 biological replicates. The boxplot indicates the median (middle line) and the third and first quartiles (box); the whiskers show the 1.5× IQR above and below the box. (**b**) MA plots depicting changes in nascent transcription (SLAM-seq) at the indicated times after dTAG-13 treatment in INTS11–FKBP cells. n = 3 biological replicates. CPM, counts per million mapped reads. (**c**) H3K4me3 and INTS11 occupancy in mES cells. The enrichments were plotted over the transcription start sites (TSS ± 2 kb) of protein-coding genes. Rows are sorted by decreasing H3K4me3 ChIP–seq occupancy in the WT mES cells. (**d**) An IGV snapshot comparing ChIP–seq signals in control and dTAG-13-treated INTS11–FKBP cells.

Extended Data Fig. 11 A Model for the roles of H3K4me3 in transcription regulation.

H3K4me3 facilitates the recruitment of factors regulating the release of paused RNAPII at the +1 nucleosome. The H3K4me3 at promoter regions is highly dynamic, and it is maintained by an equilibrium between SET1/COMPASS complexes and KDM5 demethylases at highly transcribed genes. The rapid turnover of H3K4me3 ensures that the pausing step is a highly regulated process by Integrator Complex Subunit 11 (INTS11), where an increase in H3K4me3 leads to a decrease in RNAPII pausing and acute depletion leads to an increase RNAPII pausing.

Supplementary information

Supplementary Information

Supplementary Figs. 1–5.

Reporting Summary

Peer Review File

Supplementary Tables 1–4

Supplementary Table 1: primers and oligos used for gene editing, ChIP–qPCR and RT–qPCR analysis in this study. Supplementary Table 2: antibodies used in this study. Supplementary Table 3: MS identification of proteins associated with initiating RNAPII. Supplementary Table 4: APEX2 MS identification of RNAPII neighbourhood interactions

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the

material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Wang, H., Fan, Z., Shliaha, P.V. *et al.* H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature* **615**, 339–348 (2023). https://doi.org/10.1038/s41586-023-05780-8

- Received: 07 December 2021
- Accepted: 02 February 2023
- Published: 01 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05780-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05780-8

- Matters Arising
- Open Access
- Published: 08 March 2023

Multivariate BWAS can be replicable with moderate sample sizes

- Tamas Spisak ,
- Ulrike Bingel &
- Tor D. Wager

Nature **volume 615**, pages E4–E7 (2023)

- 1484 Accesses
- 13 Altmetric
- Metrics details

Matters Arising to this article was published on 08 March 2023

The Original Article was published on 16 March 2022

arising from: S. Marek et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* https://doi.org/10.1038/s41586-022-04492-9 (2022)

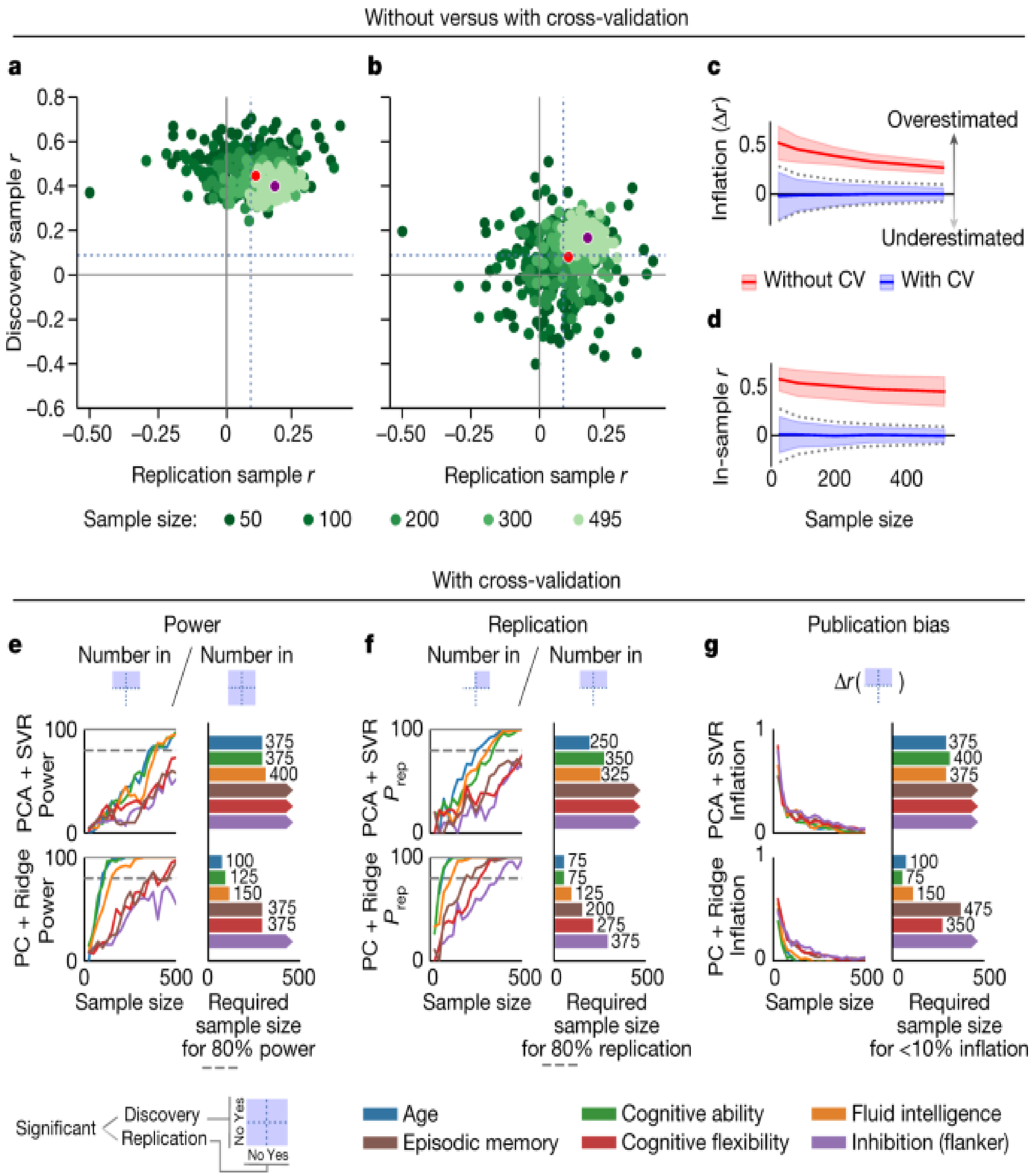
Brain-wide association studies (BWAS)—which correlate individual differences in phenotypic traits with measures of brain structure and function

—have become a dominant method for linking mind and brain over the past 30 years. Univariate BWAS typically test tens to hundreds of thousands of brain voxels individually, whereas multivariate BWAS integrate signals across brain regions into a predictive model. Numerous problems have been raised with univariate BWAS, including a lack of power and reliability and an inability to account for pattern-level information embedded in distributed neural circuits[1,2,3,4]. Multivariate predictive models address many of these concerns, and offer substantial promise for delivering brain-based measures of behavioural and clinical states and traits[2,3].

In their recent paper[4], Marek et al. evaluated the effects of sample size on univariate and multivariate BWAS in three large-scale neuroimaging datasets and came to the general conclusion that "BWAS reproducibility requires samples with thousands of individuals". We applaud their comprehensive analysis, and we agree that (1) large samples are needed when conducting univariate BWAS and (2) multivariate BWAS reveal substantially larger effects and are therefore more highly powered.

Marek et al.[4] find that multivariate BWAS provide inflated in-sample associations that often cannot be replicated (that is, are underpowered) unless thousands of participants are included. This implies that effect-size estimates from the discovery sample are necessarily inflated. However, we distinguish between the effect-size estimation method (in-sample versus cross-validated) and the sample (discovery versus replication), and show that, with appropriate cross-validation, the in-sample inflation that Marek et al.[4] report in the discovery sample can be entirely eliminated. With additional analyses, we demonstrate that multivariate BWAS effects in high-quality datasets can be replicable with substantially smaller sample sizes in some cases. Specifically, applying a standard multivariate prediction algorithm to functional connectivity in the Human Connectome Project yielded replicable effects with sample sizes of 75–500 for 5 of 6 phenotypes tested (Fig. 1).

Fig. 1: Examples of multivariate BWAS providing unbiased effect sizes and high replicability with low to moderate sample sizes.



a, Discovery sample effects in multivariate BWAS are inflated only if estimates are obtained without cross-validation (CV). **b**, Cross-validation fully eliminates in-sample effect-size inflation and, as a consequence, provides higher replicability. Data are from the Human Connectome Project (HCP1200, PTN release, $n = 1{,}003$). Each point in **a** and **b** corresponds to one bootstrap subsample, as in figure 4b of Marek et al.[4]. The dotted lines denote the threshold for $P = 0.05$ with $n = 495$. Mean multivariate brain–

behavioural phenotype associations across 100 bootstrap samples at $n = 200$ and for the full sample are denoted by red and purple dots. **c**, The inflation of in-sample effect size obtained without cross-validation (red) is reduced, but does not disappear, at higher sample sizes. Conversely, cross-validated estimates (blue) are slightly pessimistic with low sample sizes and become quickly unbiased as sample size is increased. **d**, Without cross-validation, in-sample effect-size estimates are non-zero ($r \approx 0.5$, red), even when predicting permuted outcome data. Cross-validation eliminates systematic bias across all sample sizes (blue). The dashed lines in **c** and **d** denote 95% parametric confidence intervals, and the shaded areas denote bootstrap- and permutation-based confidence intervals. **e**,**f**, Cross-validated analysis reveals that sufficient in-sample power (**e**) and out-of-sample replication probability (P_{rep}) (**f**) can be achieved for a variety of phenotypes at low or moderate sample sizes. 80% power and P_{rep} are achievable in <500 participants for 3 out of 6 phenotypes (coloured bars) using the prediction algorithm of Marek et al.[4] (**e** and **f** (top), the sample size required for 80% power or P_{rep} is shown). The remaining three phenotypes require sample sizes of >500 (bars with arrows). Power and P_{rep} can be substantially improved with a ridge regression-based model recommended in some comparison studies[10,11] (**e** and **f** (bottom), with 80% power and P_{rep} with sample sizes as low as $n = 100$ and $n = 75$, respectively, when predicting cognitive ability, and sample sizes between 75 and 375 for other investigated variables (fluid intelligence, episodic memory and cognitive flexibility), except inhibition assessed with the flanker task, which replicated with $n = 375$ but did not reach 80% power with $n = 500$. **g**, We estimated interactions between sample size and publication bias by computing effect size inflation ($r_{discovery} - r_{replication}$) only for those bootstrap cases in which prediction performance was significant ($P > 0.05$) in the replication sample. Our analysis shows that the effect-size inflation due to publication bias is modest (<10%) with fewer than 500 participants for half of the phenotypes using the model from Marek et al.[4] and all phenotypes but the flanker using the ridge model. The blue squares show conditional relationships assessed to derive metrics in **e**,**f** and **g** with reference to **b**. The top and bottom squares indicate positive and negative results in the discovery sample, respectively. The left and right squares indicate negative and positive results in the replication sample. The

blue squares indicate how these conditions were applied to derive the metrics.

These analyses are limited to a selected number of phenotypes in a relatively high-quality dataset (measured in a young adult population with a single scanner) and should not be overgeneralized. However, they highlight that the key determinant of sample size requirements is the true effect size of the brain–phenotype relationship and that, with proper internal validation, appropriate effect-size estimates and sufficiently large effects for moderately sized studies are possible.

Marek et al.[4] evaluate in-sample effect-size inflation in multivariate BWAS by training various multivariate models in a ‘discovery sample’ and comparing the in-sample effect sizes (prediction–outcome correlation, r) estimated from the training sample to the performance in an independent replication sample. On the basis of a bootstrap analysis, with variously sized pairs of samples drawn randomly from the Adolescent Brain Cognitive Development study, the authors report a severe effect-size inflation of $\Delta r = -0.29$ (average difference between the in-sample effect sizes in the discovery sample and the out-of-sample effect sizes in the replication sample) and conclude that “[e]ven at the largest sample sizes ($n \approx 2{,}000$), multivariate in-sample associations remained inflated on average”.

The issue with claims of inflation is that the in-sample effect size estimates of Marek et al.[4] were based on training multivariate models on the entire discovery sample, without cross-validation or other internal validation (as confirmed by inspection of the code and discussion with the authors). Such in-sample correlations are not valid effect-size estimates, as they produce a well-known overfitting bias that increases with model complexity[5]. Standard practice in machine learning is to evaluate model accuracy (and other performance metrics) on data independent of those used for training. In line with current recommendations for multivariate brain–behaviour analyses[6,7], this is typically performed using internal cross-validation (for example, *k*-fold) to estimate unbiased effect sizes in a discovery sample, and (less commonly) further validating significant cross-validated effects in held-out or subsequently acquired replication samples[2,5].

Using cross-validation to estimate discovery-sample effects impacts the pool of studies selected for replication attempts, the degree of effect-size attenuation in replication samples, and the sample size needed for effective replication and mitigation of publication bias. To demonstrate this and provide quantitative estimates of sample size requirements in multivariate BWAS, we analysed functional connectivity data from the Human Connectome Project[8] (one of the datasets in Marek et al.[4]) using cross-validation to estimate discovery-sample effect sizes. As shown in Fig. 1a–d, cross-validated discovery effect-size estimates are unbiased (that is, not inflated on average), irrespective of the sample size and the magnitude of the effect. As expected, even with cross-validation, smaller sample sizes resulted in lower power (Fig. 1e) and increased variability in effect-size estimates across samples (Fig. 1c). Such variability is undesirable because it reduces the probability of independent replication (Fig. 1f). Moreover, selection biases—most notably, publication bias—can capitalize on such variability to inflate effect sizes in the literature (Fig. 1g). Although these effects of using small sample sizes are undesirable, they do not invalidate the use of multivariate BWAS in small samples, and publication biases can be mitigated by practices that, like internal cross-validation, are quickly becoming standards in the field[2,5]. These include preregistration, registered reports, reporting confidence intervals and the use of hold-out samples tested only once on a single, optimized model to avoid overfitting.

Given these considerations, we wondered how many participants are generally required for multivariate BWAS. The answer to this question depends on the reliability of both phenotypic and brain measures, the size of the effects linking them, the algorithm and model-selection steps used and the use cases for the resulting brain measures. For example, multivariate models trained on as few as 20 participants[9] can have high reliability (ICC = 0.84)[10], broad external validity and large effect sizes (Hedges $g = 2.3$)[11] in independent samples (for example, more than 600 participants from 20 independent studies in ref. [11]) when predicting behavioural states within-person rather than traits. In this case, the benefit of large samples is primarily in accurately estimating local brain weights (model parameters)[12] rather than increasing out-of-sample accuracy. Here we performed functional connectivity-based multivariate BWAS with cognitive ability (the phenotype shown in figure 4 of Marek et al.[4]) and five other cognition-

related example phenotypes selected at random and demonstrate that, even when predicting trait-level phenotypes, as Marek et al.[4] did, sample sizes of 75–500 are sufficient in five out of six of cases that we tested (or three out of six cases using the prediction algorithm of Marek et al.[4]) to achieve high statistical power and replicability (for example, 80%) and to mitigate effect size inflation due to publication bias.

The basis for these estimates is shown in Fig. 1e–g. Using cross-validated discovery sample effect-size estimates, the multivariate BWAS model of Marek et al.[4]—principal-component-based reduction of bivariate connectivity followed by support vector regression (PCA + SVR)—showed 80% in-sample power and 80% out-of-sample replication probability (P_{rep}) at $n < 500$ for three out of six phenotypes that we examined (age, cognitive ability and fluid intelligence). However, this model has been shown to be disadvantageous in some comparison studies[12,13]. We therefore performed the same power and sample-size calculations for a multivariate BWAS using another approach—ridge regression on partial correlation matrices with a default shrinkage parameter of 1 (PC + ridge; Supplementary Methods). Although this approach is still probably suboptimal[12,13] (we avoided testing other models to avoid overfitting), it substantially improved the power (Fig. 1e (bottom)), independent replication probability ([P_{rep}]; Fig. 1f (bottom)) and resistance to inflation due to publication bias (Fig. 1g (bottom)). Eighty per cent power and P_{rep} were achieved at sample sizes from 75 to 150 for age (included as a reference variable), cognitive ability and fluid intelligence, and sample sizes < 400 for all phenotypes except for inhibition measured by the flanker task (a measure that is known to have low reliability[14]).

Our results highlight, that the key determinant of sample size requirements is the true effect size of the brain–phenotype relationship, which subsumes the amount, quality, homogeneity and reliability of both brain and phenotypic measures, and the degree to which a particular brain measure is relevant to a particular phenotype. Effect sizes will probably vary widely across studies; for example, cortical thickness can also reliably predict 4 out of the 6 investigated phenotypes with $n < 500$, although with smaller effect sizes on average (functional connectivity, mean $r = 0.2$; cortical thickness, mean $r = 0.1$; Supplementary Fig. 2). Although our results were derived from a

relatively high-quality dataset and used an algorithm expected to yield larger effect sizes than that of Marek et al.[4], they are in agreement with analytical calculations showing that BWAS that explain more than 1% of the phenotype's variance can be replicable with sample sizes below 1,000 (Supplementary Methods). For example, a model that explains $r^2 = 0.01$ (1% of variance) achieves 80% power in a prospective replication with $n = 801$, and $r^2 = 0.02$ achieves 80% power with $n = 399$ (ref. [15]).

These quantitative differences in required sample size could translate into large, qualitative differences in the types of neuroimaging studies considered viable in future efforts. There is a necessary trade-off between the innovativeness of a task, measure or method, and the extent to which it has been validated. Existing large-scale neuroimaging studies ($n > 1{,}000$) have selected well-validated tasks and imaging measures over new, exploratory ones, and few have attempted to characterize rare populations. Requiring sample sizes that are larger than necessary for the discovery of new effects could stifle innovation.

We agree with Marek et al.[4] that small-sample studies are important for understanding the brain bases of tasks and mental states[9,10,11], and for prototyping new tasks and measures. Furthermore, several current trends may further increase the viability of small-sample multivariate BWAS, including (1) new phenotypes, (2) feature-learning methods and algorithms with larger effect sizes[13], (3) models that target within-person variation in symptoms and behaviour to improve between-person predictions[2] and (4) hybrid strategies for improving prediction like meta-matching[16]. All of these have the potential to improve reliability and effect sizes, but whether they do remains to be seen.

Finally, as both Marek et al.[4] and our analyses show, very small effects will suffer from limited power, replicability and predictive utility even with sample sizes in the thousands (Fig. 1). We argue that the field should focus on discovering phenotypes and brain measures with large effect sizes. Efficient discovery entails casting a wide net in smaller studies using rigorous, unbiased methods and scaling up promising findings to larger samples[2]. There are substantial challenges ahead, including establishing broad generalizability across contexts, equity across subpopulations, and

models with high neuroscientific validity and interpretability[17,18]. Addressing these challenges will require innovative new methods and measures.

Reporting summary

Further information on experimental design is available in the Nature Portfolio Reporting Summary linked to this Article.

Data availability

Analysis is based on preprocessed data provided by the Human Connectome Project, WU-Minn Consortium (principal investigators: D. Van Essen and K. Ugurbil; 1U54MH091657) funded by the 16 NIH institutes and centres that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. All data used in the present study are available for download from the Human Connectome Project (www.humanconnectome.org). Users must agree to data use terms for the HCP before being allowed access to the data and ConnectomeDB; details are provided online (https://www.humanconnectome.org/study/hcp-young-adult/data-use-terms).

Code availability

All analysis code used in the current study is available at GitHub (release v.1.0; https://github.com/spisakt/BWAS_comment).

References

1. Button, K. S. et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14**, 365–376 (2013).

2. Woo, C.-W. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: Brain models in translational neuroimaging. *Nat. Neurosci.* **20**, 365–377 (2017).

3. Bzdok, D., Varoquaux, G. & Steyerberg, E. W. Prediction, not association, paves the road to precision medicine. *JAMA Psychiatry* **78**, 127–128 (2021).

4. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).

5. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* **77**, 534–540 (2020).

6. Genon, S., Eickhoff, S. B. & Kharabian, S. Linking interindividual variability in brain structure to behaviour. *Nat. Rev. Neurosci.* **23**, 307–318 (2022).

7. Rosenberg, M. D. & Finn, E. S. How to establish robust brain–behavior relationships without thousands of individuals. *Nat. Neurosci.* **25**, 835–837 (2022).

8. Van Essen, D. C. et al. The WU-Minn human connectome project: an overview. *Neuroimage*, **80**, 62–79 (2013).

9. Wager, T. D. et al. An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* **368**, 1388–1397 (2013).

10. Han, X. et al. Effect sizes and test-retest reliability of the fMRI-based neurologic pain signature. *Neuroimage* **247**, 118844 (2022).

11. Zunhammer, M., Bingel, U. & Wager, T. D. Placebo effects on the neurologic pain signature. *JAMA Neurol.* **75**, 1321–1330 (2018).

12. Tian, Y. & Zalesky, A. Machine learning prediction of cognition from functional connectivity: are feature weights reliable? *Neuroimage* **245**, 118648 (2021).

13. Pervaiz, U., Vidaurre, D., Woolrich, M. W. & Smith, S. M. Optimising network modelling methods for fMRI. *Neuroimage* **211**, 116604 (2020).

14. Hedge, C., Powell, G. & Sumner, P. The reliability paradox: why robust cognitive tasks do not produce reliable individual differences. *Behav. Res. Methods* **50**, 1166–1186 (2018).

15. Killeen, P. R. Predict, control, and replicate to understand: how statistics can foster the fundamental goals of science. *Perspect. Behav. Sci.* **42**, 109–132 (2018).

16. He, T. et al. Meta-matching as a simple framework to translate phenotypic predictive models from big to small data. *Nat. Neurosci.* **25**, 795–804 (2022).

17. Wu, J. et al. Cross-cohort replicability and generalizability of connectivity-based psychometric prediction patterns. *Neuroimage* **262**, 119569 (2022).

18. Li, J. et al. Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity. *Sci. Adv.* **8**, eabj1812 (2022).

Acknowledgements

We thank S. Marek et al. for sharing the analysis code and for the discussions in relation to our commentary. The work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; projects ‘TRR289 - Treatment Expectation’, ID 422744262 and ‘SFB1280 - Extinction Learning’, ID 316803389), R01 MH076136 and R01 EB026549.

Author information

Authors and Affiliations

1. Institute of Diagnostic and Interventional Radiology and Neuroradiology, University Medicine Essen, Essen, Germany

   Tamas Spisak

2. Center for Translational Neuro- and Behavioral Sciences, Department of Neurology, University Medicine Essen, Essen, Germany

   Tamas Spisak & Ulrike Bingel

3. Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

   Tor D. Wager

Contributions

Conception and data analysis: T.S. Manuscript writing and revision: T.S., U.B. and T.D.W.

Corresponding author

Correspondence to Tamas Spisak.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary Information

Supplementary Figs. 1 and 2, Supplementary Methods and Supplementary References.

[Reporting Summary]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Spisak, T., Bingel, U. & Wager, T.D. Multivariate BWAS can be replicable with moderate sample sizes. *Nature* **615**, E4–E7 (2023). https://doi.org/10.1038/s41586-023-05745-x

- Received: 10 April 2022
- Accepted: 19 January 2023
- Published: 08 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05745-x

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05745-x

- Matters Arising
- Open Access
- Published: 08 March 2023

Reply to: Multivariate BWAS can be replicable with moderate sample sizes

- Brenden Tervo-Clemmens ,
- Scott Marek ,
- Roselyne J. Chauvin,
- Andrew N. Van,
- Benjamin P. Kay,
- Timothy O. Laumann,
- Wesley K. Thompson,
- Thomas E. Nichols,
- B. T. Thomas Yeo,
- Deanna M. Barch,
- Beatriz Luna,
- Damien A. Fair &
- …
- Nico U. F. Dosenbach

Show authors

Nature **volume 615**, pages E8–E12 (2023)

- 1649 Accesses
- 42 Altmetric
- Metrics details

The Original Article was published on 08 March 2023

replying to: T. Spisak et al. *Nature* https://doi.org/10.1038/s41586-023-05745-x (2023)

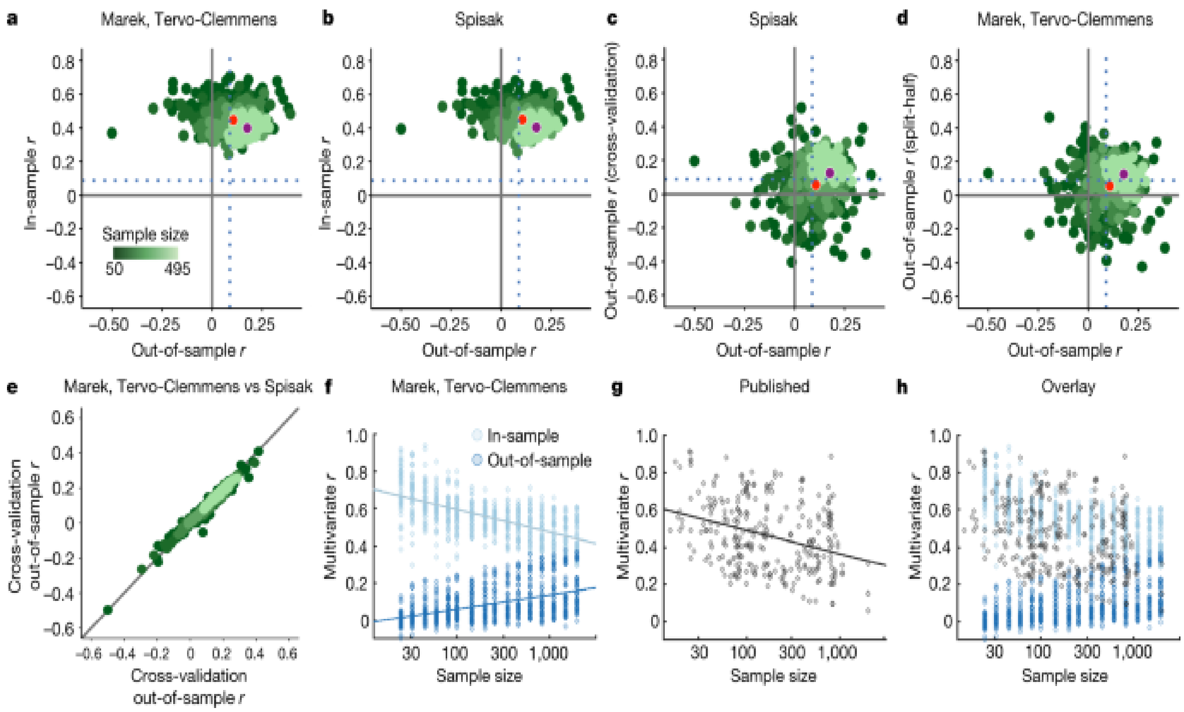
In our previous study[1], we documented the effect of sample size on the reproducibility of brain-wide association studies (BWAS) that aim to cross-sectionally relate individual differences in human brain structure (cortical thickness) or function (resting-state functional connectivity (RSFC)) to cognitive or mental health phenotypes. Applying univariate and multivariate methods (for example, support vector regression (SVR)) to three large-scale neuroimaging datasets (total $n \approx 50{,}000$), we found that overall BWAS reproducibility was low for $n < 1{,}000$, due to smaller than expected effect sizes. When samples and true effects are small, sampling variability, and/or overfitting can generate 'statistically significant' associations that are likely to be reported due to publication bias, but are not reproducible[2,3,4,5], and we therefore suggested that BWAS should build on recent precedents[6,7] and continue to aim for samples in the thousands. In the accompanying Comment, Spisak et al.[8] agree that larger BWAS are better[5,9], but argue that "multivariate BWAS effects in high-quality datasets can be replicable with substantially smaller sample sizes in some cases" ($n = 75$–500); this suggestion is made on the basis of analyses of a selected subset of multivariate cognition/RSFC associations with larger effect sizes, using their preferred method (ridge regression with partial correlations) in a demographically more homogeneous, single-site/scanner sample (Human Connectome Project (HCP), $n = 1{,}200$, aged 22–35 years).

There is no disagreement that a minority of BWAS effects can replicate in smaller samples, as shown with our original methods[1]). Using the exact methodology (including cross-validation) and code of Spisak et al.[8] to repeat 64 multivariate BWAS in the 21-site, larger and more diverse Adolescent Brain Cognitive Development Study (ABCD, $n = 11{,}874$, aged 9–11 years), we found that 31% replicated at $n = 1{,}000$, dropping to 14% at $n = 500$ and none at $n = 75$. Contrary to the claims of Spisak et al.[8], replication failure was the outcome in most cases when applied to this larger, more diverse dataset. Basing general BWAS sample size recommendations on the largest

effects has at least two fundamental flaws: (1) failing to detect other true effects (for example, reducing the sample size from $n = 1,000$ to $n = 500$ leads to a 55% false-negative rate), therefore restricting BWAS scope, and (2) inflation of reported effects[3,10,11,12]. Thus, regardless of the method, associations based on small samples can remain distorted and lack generalizability until confirmed in large, diverse, independent samples.

We always test for BWAS replication with null models (using permutation tests) of out-of-sample estimates to ensure that our reported reproducibility is unaffected by in-sample overfitting. Nonetheless, Spisak et al.[8] argue against plotting inflated in-sample estimates[1,10] on the *y* axis, and out-of-sample values on the *x* axis, as we did (Fig. 1a). Instead, they propose plotting cross-validated associations from an initial, discovery sample (Fig. 1b (*y* axis)) against split-half out-of-sample associations (*x* axis). However, cross-validation—just like split-half validation—estimates out-of-sample, and not in-sample, effect sizes[13]. The in-sample associations[1,10] for the method of Spisak et al.[8] (Fig. 1b), that is, from data in the sample used to develop the model, show the same degree of overfitting (Fig. 1a versus Fig. 1b). The plot of Spisak et al.[8] (Fig. 1c) simply adds an additional out-of-sample test (cross-validation before split half), and therefore demonstrates the close correspondence between two different methods for out-of-sample effect estimation[14]. Analogously, we can replace the cross-validation step in the code of Spisak et al.[8] with split-half validation (our original out-of-sample test), obtaining split-half effects in the first half of the sample, and then comparing them to the split-half estimates from the full sample (Fig. 1d). The strong correspondences between cross-validation followed by split-half (Spisak et al. method[8]; Fig. 1c) and repeated split-half validation (Fig. 1d) are guaranteed by plotting out-of-sample estimates (from the same dataset) against one another. Here, plotting cross-validated discovery sample estimates on the *y* axis (Fig. 1c,d) provides no additional information beyond the *x* axis out-of-sample values. The critically important out-of-sample predictions, required for reporting multivariate results[1], generated using the method of Spisak et al.[8] and our method are nearly identical (Fig. 1e).

Fig. 1: In-sample versus out-of-sample effect estimates in multivariate BWAS.



a–**e**, Methods comparison between our previous study[1] (split-half) and Spisak et al.[8] (cross-validation followed by split-half). 'Marek, Tervo-Clemmens' and 'Spisak' refer to the methodolgies described in ref. [1] and ref. [8], respectively. For **a**–**e**, HCP 1200 Release (full correlation) data were used to predict age-adjusted total cognitive ability. Analysis code and visualizations (*x*,*y* scaling; colours) are the same as in Spisak et al.[8]. The *x* axes in **a**–**e** always display the split-half out-of-sample effect estimates from the second (replication) half of the data (correlation between true scores and predicted scores; as in Spisak et al.[8] and in our previous study[1]; Supplementary Methods). **a**, In-sample (training correlation; *y* axis) as a function of out-of-sample associations (plot convention in our previous study[1]). **b**, Matched comparison of the true in-sample association (training correlations, mean across folds; *y* axis) in the method proposed by Spisak et al.[8]. **c**, The proposed correction by Spisak et al.[8] that inserts an additional cross-validation step to evaluate the first half of data, which by definition makes this an out-of-sample association (*y* axis). **d**, Replacing the cross-validation step from Spisak et al.[8] with a split-half validation provides a different (compared with **c**) out-of-sample association of the first half of the total data (that is, each of the first stage split halves is one-quarter of the total data; *y* axis). The appropriate and direct comparison of in-sample

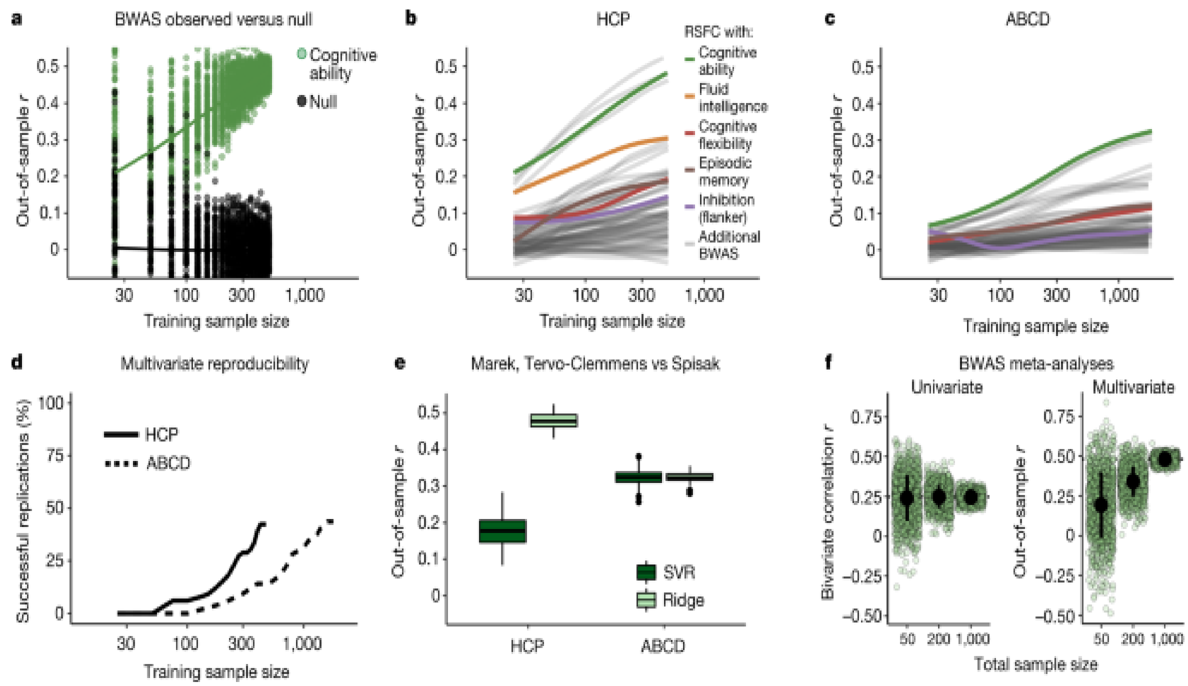
associations between Spisak et al.[8] and our previous study[1] is comparing **b** to **a**, rather than **c** to **a**. The Spisak et al. method[8] (cross-validation followed by split-half validation) does not reduce in-sample overfitting (**b**) but, instead, adds an additional out-of-sample evaluation (**c**), which is nearly identical to split-half validation twice in a row (**d**), and makes it clear why the out-of-sample performance of these two methods is likewise nearly identical. **e**, Correspondence between out-of-sample associations (to the left-out half) from the additional cross-validation step proposed by Spisak et al.[8] (mean across folds; *y* axis) and the original split-half validation from our previous study[1] (*x* axis). The identity line is shown in black. **f**, In-sample (*r*; light blue) and out-of-sample (*r*; dark blue) associations as a function of sample size. Data are from figure 4a–d of ref. [1]. **g**, Published literature review of multivariate *r* (*y* axis) as a function of sample size (data from ref. [15]) displayed with permission. For **f** and **g**, best fit lines are displayed in $\log_{10}$ space. **h**, Overlap of **f** and **g**.

As Spisak et al.[8] highlight, cross-validation of some type is considered to be standard practice[10], and yet the distribution of out-of-sample associations (Fig. 1f (dark blue)) does not match published multivariate BWAS results (Fig. 1g), which have largely ranged from $r = 0.25$ to 0.9, decreasing with increasing sample size[10,15,16]. Instead, published effects more closely follow the distribution of in-sample associations (Fig. 1h). This observation suggests that, in addition to small samples, structural problems in academic research (for example, non-representative samples, publication bias, misuse of cross-validation and unintended overfitting) have contributed to the publication of inflated effects[12,17,18]. A recent biomarker challenge[5] showed that cross-validation results continued to improve with the amount of time researchers spent with the data, and the models with the best cross-validation results performed worse on never-seen held-back data. Thus, cross-validation alone has proven to be insufficient and must be combined with the increased generalizability of large, diverse datasets and independent out-of-sample evaluation in new, never before seen data[5,10].

The use of additional cross-validation in the discovery sample by Spisak et al.[8] does not affect out-of-sample prediction accuracies (Fig. 1e). However, by using partial correlations and ridge regression on HCP data, they were

able to generate higher out-of-sample prediction accuracies than our original results in ABCD (Fig. 2a). The five variables they selected are strongly correlated[19] cognitive measures from the NIH Toolbox (mean $r = 0.37$; compare with the correlation strength for height versus weight $r = 0.44$)[20] and age (not a complex behavioural phenotype), unrepresentative of BWAS as a whole (Fig. 2b (colour versus grey lines)). As the HCP is the relatively smallest and most homogeneous dataset, we applied the exact method and code of Spisak et al.[8] to the ABCD data (Fig. 2c and Supplementary Table 2). At $n = 1,000$ (training; $n = 2,000$ total), only 31% of BWAS (44% RSFC, 19% cortical thickness) were replicable (Fig. 2d; defined as in Spisak et al.[8]; Supplementary Information). Expanding BWAS scope beyond broad cognitive abilities towards complex mental health outcomes therefore requires $n > 1,000$ (Fig. 2b–d). The absolute largest BWAS (cognitive ability: RSFC, green) reached replicability only using $n = 400$ ($n = 200$ train; $n = 200$ test) with an approximate 40% decrease in out-of-sample prediction accuracies from HCP to ABCD (Fig. 2e (lighter green, left versus right)). The methods of Spisak et al.[8] and our previous study[1] returned equivalent out-of-sample reproducibility for this BWAS (cognitive ability: RSFC) in the larger, more diverse ABCD data (Fig. 2e (right, dark versus light green)). Thus, the smaller sample sizes (Fig. 2b,c) that are required for out-of-sample reproducibility (Fig. 2e) reported by Spisak et al.[8] in the HCP data did not generalize to the larger ABCD dataset. See also our previous study[1] for a broader discussion of convergent evidence across HCP and ABCD datasets.

Fig. 2: BWAS reproducibility, scope and prediction accuracy using the method of Spisak et al.



a, Example bootstrapped BWAS of total cognitive ability (green) and null distribution (black) (*y* axis), as a function of sample size (*x* axis) from the suggested method of Spisak et al.[8] (RSFC by partial correlation; prediction by ridge regression) in the HCP dataset ($n = 1{,}200$, 1 site, 1 scanner, 60 min RSFC/participant, 76% white). Sample sizes were $\log_{10}$-transformed for visualization. **b**, Out-of-sample correlation (between true scores and predicted scores) from ridge regression (*y* axis; code from Spisak et al.[8]) as a function of training sample size (*x* axis, $\log_{10}$ scaling) for 33 cognitive and mental health phenotypes (Supplementary Information) in the HCP dataset. Each line displays a smoothed fit estimate (through penalized splines in general additive models) for a brain (RSFC (partial correlations, as proposed by Spisak et al.[8]), cortical thickness) phenotype pair (66 total) that has 100 bootstrapped iterations from sample sizes of 25 to 500 (inclusive) in increments of 25 (20 total bins). Sample sizes were $\log_{10}$-transformed (for visualization) before general additive model fitting. **c**, The same as in **b**, but in the ABCD dataset ($n = 11{,}874$, 21 sites, 3 scanner manufacturers, 20 min RSFC/participant, 56% white) using 32 cognitive and mental health phenotypes at sample sizes of 25, 50, 75 and from 100 to 1,900 (inclusive) in increments of 100 (22 total bins). **d**, The percentage of brain–phenotype pairs (BWAS) from **b** and **c** with significant replication on the basis of the method of Spisak et al.[8] (Supplementary Information). **e**, Comparison of our

original method in our previous study[1] and the method proposed by Spisak et al.[8] at the full split-half sample size of HCP (left) and ABCD (right). Out-of-sample correlations (RSFC with total cognitive ability, *y* axis) for the method used in our previous study[1] (dark green; RSFC by correlation, PCA, SVR) and by Spisak et al.[8] (light green; RSFC by partial correlation, ridge regression). Repeating the method proposed by Spisak et al.[8] in ABCD (right) and comparing this to the method used in our previous study[1] results in a very similar out-of-sample *r*. **f**, Simulated individual studies (light green circles; $n = 1{,}000$ per sample size) and meta-analytic estimates (black dot, ±1 s.d.) using the method of Spisak et al.[8] (partial correlations in the HCP dataset) for the largest univariate association (left; *y* axis, bivariate correlation) and multivariate association (right; *y* axis, out-of-sample correlation) for total cognitive ability versus RSFC, as a function of total sample size (*x* axis; bivariate correlation for sample sizes of 50, 200 and 1,000, and multivariate sum of train and test samples, each 25, 100 and 500). For univariate approaches, studies of any sample size, when appropriately aggregated to a large total sample size, can correctly estimate the true effect size. However, for multivariate approaches, even when aggregating across 1,000 independent studies, studies with a small sample size produce prediction accuracies that are downwardly biased relative to large sample studies, highlighting the need for large samples in multivariate analyses.

Notably, the objections of Spisak et al.[8] raise additional reasons to stop the use of smaller samples in BWAS that were not highlighted in our original article. Multivariate BWAS prediction accuracies—absent overfitting—are systematically suppressed in smaller samples[5,9,21], as prediction accuracy scales with increasing sample size[1,9]. Thus, the claim that “cross-validated discovery effect-size estimates are unbiased” does not account for out-of-dataset generalizability and downward bias. In principle, if unintended overfitting and publication bias could be fully eliminated, meta-analyses of small-sample univariate BWAS would return the correct association strengths (Fig. 2f (left)). However, meta-analyses of small multivariate BWAS would always be downwardly biased (Fig. 2f (right)). If we are interested in maximizing prediction accuracy, essential for clinical implementation of BWAS[22], large samples and advancements in imaging and phenotypic measurements[1] are necessary.

Repeatedly subsampling the same dataset, as Spisak et al.[8] and we have done, overestimates reproducibility compared with testing on a truly new, diverse dataset. Just as in genomics[23], BWAS generalization failures have been highlighted[5,24]. For example, BWAS models trained on white Americans transferred poorly to African Americans and vice versa (within dataset)[24]. Historically, BWAS samples have lacked diversity, neglecting marginalized and under-represented minorities[25]. Large studies with more diverse samples and data aggregation efforts can improve BWAS generalizability and reduce scientific biases contributing to massive health inequities[26,27].

Spisak et al.[8] worry that "[r]equiring sample sizes that are larger than necessary for the discovery of new effects could stifle innovation". We appreciate the concern that rarer populations may never be investigated with BWAS. Yet, there are many non-BWAS brain–behaviour study designs (fMRI ≠ BWAS) focused on within-patient effects, repeated-sampling and signal-to-noise-ratio improvements that have proven fruitful down to $n = 1$ (ref. [28]). By contrast, the strength of multivariate BWAS lies in leveraging large cross-sectional samples to investigate population-level questions. Sample size requirements should be based on expected effect sizes and real-world impact, and not resource availability. Through large-scale collaboration and clear standards on data sharing, GWAS has reached sample sizes in the millions[29,30,31], pushing genomics towards new horizons. Similarly, BWAS analyses of the future will not be limited to statistical replication of the same few strongest effects in small homogeneous populations, but also have broad scope, maximum prediction accuracy and excellent generalizability.

Reporting summary

Further information on experimental design is available in the Nature Portfolio Reporting Summary linked to this Article.

Data availability

Participant-level data from all datasets (ABCD and HCP) are openly available pursuant to individual, consortium-level data access rules. The ABCD data repository grows and changes over time (https://nda.nih.gov/abcd). The ABCD data used in this report came from ABCD collection 3165 and the Annual Release 2.0 (https://doi.org/10.15154/1503209). Data were provided, in part, by the HCP, WU-Minn Consortium (principal investigators: D. Van Essen and K. Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Some data used in the present study are available for download from the HCP (www.humanconnectome.org). Users must agree to data use terms for the HCP before being allowed access to the data and ConnectomeDB, details are provided online (https://www.humanconnectome.org/study/hcp-young-adult/data-use-terms).

Code availability

Manuscript analysis code specific to this study is available at GitHub (https://gitlab.com/DosenbachGreene/bwas_response). Code for processing ABCD data is provided at GitHub (https://github.com/DCAN-Labs/abcd-hcp-pipeline). MRI data analysis code is provided at GitHub (https://github.com/ABCD-STUDY/nda-abcd-collection-3165). FIRMM software is available online (https://firmm.readthedocs.io/en/latest/release_notes/). The ABCD Study used v.3.0.14.

References

1. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).

2. Schönbrodt, F. D. & Perugini, M. At what sample size do correlations stabilize? *J. Res. Pers.* **47**, 609–612 (2013).

3. Button, K. S. et al. Confidence and precision increase with high statistical power. *Nat. Rev. Neurosci.* **14**, 585–586 (2013).

4. Varoquaux, G. Cross-validation failure: small sample sizes lead to large error bars. *Neuroimage* **180**, 68–77 (2018).

5. Traut, N. et al. Insights from an autism imaging biomarker challenge: promises and threats to biomarker discovery. *Neuroimage* **255**, 119171 (2022).

6. Casey, B. J. et al. The Adolescent Brain Cognitive Development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).

7. Littlejohns, T. J. et al. The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. *Nat. Commun.* **11**, 2624 (2020).

8. Spisak, T., Bingel, U. & Wager, T. D. Multivariate BWAS can be replicable with moderate sample sizes. *Nature* https://doi.org/10.1038/s41586-023-05745-x (2023).

9. Schulz, M.-A., Bzdok, D., Haufe, S., Haynes, J.-D. & Ritter, K. Performance reserves in brain-imaging-based phenotype prediction. Preprint at https://doi.org/10.1101/2022.02.23.481601 (2022).

10. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* **77**, 534–540 (2020).

11. Poldrack, R. A. The costs of reproducibility. *Neuron* **101**, 11–14 (2019).

12. Button, K. S. et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14**, 365–376 (2013).

13. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proc. IJCAI 95* (ed. Mellish, C. S.) 1137–1143 (Morgan Kaufman, 1995).

14. Scheinost, D. et al. Ten simple rules for predictive modeling of individual differences in neuroimaging. *Neuroimage* **193**, 35–45 (2019).

15. Sui, J., Jiang, R., Bustillo, J. & Calhoun, V. Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. *Biol. Psychiatry* **88**, 818–828 (2020).

16. Woo, C.-W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nat. Neurosci.* **20**, 365–377 (2017).

17. Ioannidis, J. P. A. Why most discovered true associations are inflated. *Epidemiology* **19**, 640–648 (2008).

18. Pulini, A. A., Kerr, W. T., Loo, S. K. & Lenartowicz, A. Classification accuracy of neuroimaging biomarkers in attention-deficit/hyperactivity disorder: effects of sample size and circular analysis. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **4**, 108–120 (2019).

19. Funder, D. C. & Ozer, D. J. Evaluating effect size in psychological research: sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* **2**, 156–168 (2019).

20. Meyer, G. J. et al. Psychological testing and psychological assessment: a review of evidence and issues. *Am. Psychol.* **56**, 128–165 (2001).

21. He, T. et al. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *Neuroimage* **206**, 116276 (2020).

22. Leptak, C. et al. What evidence do we need for biomarker qualification? *Sci. Transl. Med.* **9**, eaal4599 (2017).

23. Weissbrod, O. et al. Leveraging fine-mapping and multipopulation training data to improve cross-population polygenic risk scores. *Nat. Genet.* **54**, 450–458 (2022).

24. Li, J. et al. Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity. *Sci. Adv.* **8**, eabj1812 (2022).

25. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83 (2010).

26. Bailey, Z. D. et al. Structural racism and health inequities in the USA: evidence and interventions. *Lancet* **389**, 1453–1463 (2017).

27. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).

28. Gratton, C., Nelson, S. M. & Gordon, E. M. Brain-behavior correlations: two paths toward reliability. *Neuron* **110**, 1446–1449 (2022).

29. Levey, D. F et al. Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nature* **24**, 954–963 (2021).

30. Muggleton, N. et al. The association between gambling and financial, social and health outcomes in big financial data. *Nat. Hum. Behav.* **5**, 319–326 (2021).

31. Yengo, L. et al. A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).

Acknowledgements

Data used in the preparation of this Article were, in part, obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 years and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, U24DA041147, U01DA041093 and U01DA041025. A full list of supporters is available online (https://abcdstudy.org/federal-partners.html). A listing of participating sites

and a complete listing of the study investigators is available online (https://abcdstudy.org/scientists/workgroups/). ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or the writing of this report. This Article reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. Data were provided, in part, by the HCP, WU-Minn Consortium (U54 MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. This work used the storage and computational resources provided by the Masonic Institute for the Developing Brain (MIDB), the Neuroimaging Genomics Data Resource (NGDR) and the Minnesota Supercomputing Institute (MSI). The NGDR is supported by the University of Minnesota Informatics Institute through the MnDRIVE initiative in coordination with the College of Liberal Arts, Medical School and College of Education and Human Development at the University of Minnesota. This work used the storage and computational resources provided by the Daenerys Neuroimaging Community Computing Resource (NCCR). The Daenerys NCCR is supported by the McDonnell Center for Systems Neuroscience at Washington University, the Intellectual and Developmental Disabilities Research Center (IDDRC; P50 HD103525) at Washington University School of Medicine and the Institute of Clinical and Translational Sciences (ICTS; UL1 TR002345) at Washington University School of Medicine. This work was supported by NIH grants MH121518 (to S.M.), NS090978 (to B.P.K.), MH129616 (to T.O.L.), 1RF1MH120025-01A1 (to W.K.T), MH080243 (to B.L.), MH067924 (to B.L.), DA041148 (to D.A.F.), DA04112 (to D.A.F.), MH115357 (to D.A.F.), MH096773 (to D.A.F. and N.U.F.D.), MH122066 (to D.A.F. and N.U.F.D.), MH121276 (to D.A.F. and N.U.F.D.), MH124567 (to D.A.F. and N.U.F.D.), NS088590 (to N.U.F.D.), and the Andrew Mellon Predoctoral Fellowship (to B.T.-C.), the Staunton Farm Foundation (to B.L.), the Lynne and Andrew Redleaf Foundation (to D.A.F.) and the Kiwanis Neuroscience Research Foundation (to N.U.F.D.).

Author information

Author notes

1. These authors contributed equally: Brenden Tervo-Clemmens, Scott Marek

2. These authors jointly supervised this work: Damien A. Fair, Nico U. F. Dosenbach

Authors and Affiliations

1. Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

   Brenden Tervo-Clemmens

2. Department of Radiology, Washington University School of Medicine, St Louis, MO, USA

   Scott Marek & Nico U. F. Dosenbach

3. Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA

   Scott Marek, Timothy O. Laumann & Deanna M. Barch

4. Department of Neurology, Washington University School of Medicine, St Louis, MO, USA

   Roselyne J. Chauvin, Andrew N. Van, Benjamin P. Kay & Nico U. F. Dosenbach

5. Department of Biomedical Engineering, Washington University in St Louis, St Louis, MO, USA

   Andrew N. Van & Nico U. F. Dosenbach

6. Division of Biostatistics, University of California San Diego, La Jolla, CA, USA

   Wesley K. Thompson

7. Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Thomas E. Nichols

8. Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

B. T. Thomas Yeo

9. Centre for Sleep and Cognition, National University of Singapore, Singapore, Singapore

B. T. Thomas Yeo

10. Centre for Translational MR Research, National University of Singapore, Singapore, Singapore

B. T. Thomas Yeo

11. N.1 Institute for Health, Institute for Digital Medicine, National University of Singapore, Singapore, Singapore

B. T. Thomas Yeo

12. Integrative Sciences and Engineering Programme, National University of Singapore, Singapore, Singapore

B. T. Thomas Yeo

13. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA

B. T. Thomas Yeo

14. Department of Psychological and Brain Sciences, Washington University in St Louis, St Louis, MO, USA

Deanna M. Barch & Nico U. F. Dosenbach

15. Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA

    Beatriz Luna

16. Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA

    Beatriz Luna

17. Masonic Institute for the Developing Brain, University of Minnesota Medical School, Minneapolis, MN, USA

    Damien A. Fair

18. Department of Pediatrics, University of Minnesota Medical School, Minneapolis, MN, USA

    Damien A. Fair

19. Institute of Child Development, University of Minnesota Medical School, Minneapolis, MN, USA

    Damien A. Fair

20. Program in Occupational Therapy, Washington University School of Medicine, St Louis, MO, USA

    Nico U. F. Dosenbach

21. Department of Pediatrics, Washington University School of Medicine, St Louis, MO, USA

    Nico U. F. Dosenbach

Contributions

Conception: B.T.-C., S.M., D.A.F. and N.U.F.D. Design: B.T.-C., S.M., R.J.C., D.A.F. and N.U.F.D. Data acquisition, analysis and interpretation: B.T.-C., S.M., R.J.C., A.V.N., B.P.K., W.K.T., T.E.N., B.T.T.Y., D.A.F. and N.U.F.D. Manuscript writing and revising: B.T.-C., S.M., R.J.C., A.V.N., B.P.K., T.O.L., W.K.T., T.E.N., B.T.T.Y., D.M.B., B.L., D.A.F. and N.U.F.D. We note that the reply author list differs from the original paper in number and in order to accurately reflect its more focused scope compared with the original work.

Corresponding authors

Correspondence to Brenden Tervo-Clemmens, Scott Marek, Damien A. Fair or Nico U. F. Dosenbach.

Ethics declarations

Competing interests

D.A.F. and N.U.F.D. have a financial interest in Turing Medical and may financially benefit if the company is successful in marketing FIRMM motion monitoring software products. A.N.V., D.A.F. and N.U.F.D. may receive royalty income based on FIRMM technology developed at Washington University School of Medicine and Oregon Health and Sciences University and licensed to Turing Medical. D.A.F. and N.U.F.D. are co-founders of Turing Medical. These potential conflicts of interest have been reviewed and are managed by Washington University School of Medicine, Oregon Health and Sciences University and the University of Minnesota.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary Information

Supplementary Methods, Supplementary Tables 1 and 2 and Supplementary References.

Reporting Summary

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Tervo-Clemmens, B., Marek, S., Chauvin, R.J. *et al.* Reply to: Multivariate BWAS can be replicable with moderate sample sizes. *Nature* **615**, E8–E12 (2023). https://doi.org/10.1038/s41586-023-05746-w

- Published: 08 March 2023
- Issue Date: 09 March 2023
- DOI: https://doi.org/10.1038/s41586-023-05746-w

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Reproducible brain-wide association studies require thousands of individuals

- Scott Marek
- Brenden Tervo-Clemmens
- Nico U. F. Dosenbach

Nature Article Open Access 16 Mar 2022

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05746-w

Collections

- **Chronic kidney disease** [08 Mar 2023]
  • Every 30 minutes, our kidneys filter all the blood in the body.
- **Japan** [08 Mar 2023]
  • Japan is the world's third largest economy, but for years now it has lagged in science and research compared to smaller economies such as Germany and the United Kingdom.



Researchers tackle chronic kidney disease

Innovative technology and a class of drugs called SGLT2 inhibitors could help more people with this common condition.

- Herb Brody

Outlook 8 Mar 2023 Nature



SGLT2 inhibitors breathe life into kidney-disease care

Researchers want to expand the use of drugs that protect the hearts and kidneys of people with chronic kidney disease.

- Amanda B. Keener

Outlook 8 Mar 2023 Nature



Stop denying kidney transplants to non-citizens

Policies that prevent undocumented immigrants in the United States from receiving a donated organ are not only cruel and unfair, but also fail the cost-effectiveness test.

- C. Elena Cervantes

Outlook 8 Mar 2023 Nature



The pursuit of dialysis equity

When a transplant is out of reach, kidney failure leaves those without access to high-quality health care with few options.

- Lauren Gravitz

Outlook 8 Mar 2023 Nature



[Confronting racial and ethnic disparities in diabetic kidney disease]

Diabetes is on the rise globally, mostly in low- and middle-income countries and among minority ethnic groups in wealthier nations. This increase is behind a surge in chronic kidney disease.

- Charles Schmidt

Outlook 8 Mar 2023 Nature

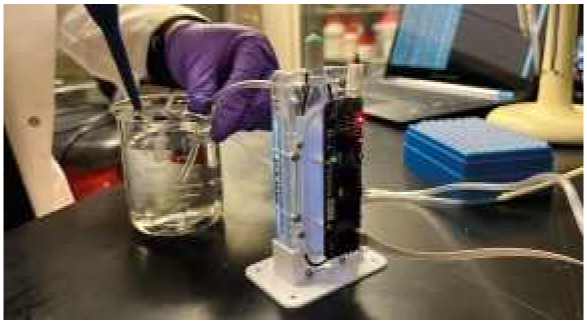


How organoids are advancing the understanding of chronic kidney disease

Although complete human kidneys grown from scratch are many years away, organoids built from pluripotent stem cells are already helping to model the condition and suggest better treatments.

- Eric Bender

Outlook 8 Mar 2023 Nature

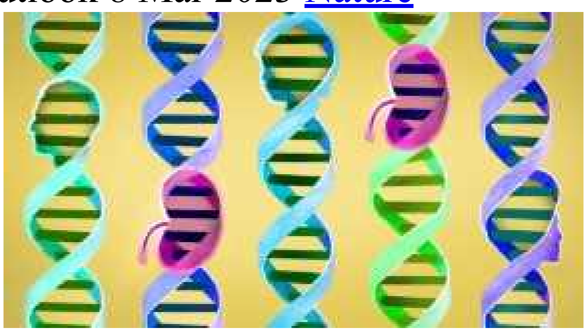


Could implantable artificial kidneys end the need for dialysis?

Highly engineered mechanical structures could radically improve the quality of life for people with chronic kidney disease.

- Neil Savage

Outlook 8 Mar 2023 Nature

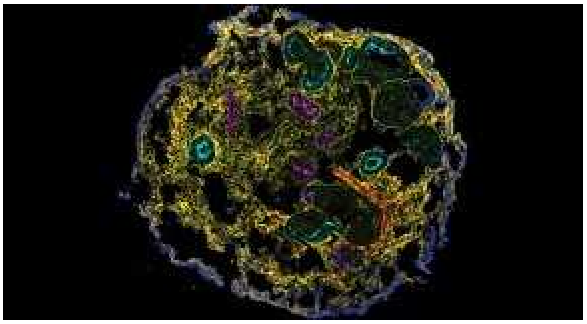


The genetic revolution transforming kidney disease

Genetic sequencing is changing the way the often deadly disorder is diagnosed, managed, treated and prevented.

- Bianca Nogrady

Outlook 8 Mar 2023 Nature



Chronic kidney disease: highlights from research

Illuminating genetic risk, disrupting fibrosis and intercepting inflammation.

- Michael Eisenstein

Outlook 8 Mar 2023 Nature



CKD is a hidden threat, but the treatment landscape is evolving

Early screening and cardio–renal drug combinations are key to preventing more deaths from chronic kidney disease.

Sponsor Feature 8 Mar 2023

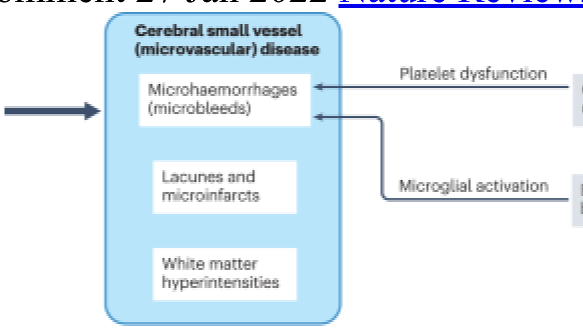
An integrated approach towards a public health perspective on chronic kidney disease

Increased awareness of chronic kidney disease among the general public is required to facilitate action to improve kidney health. An integrated approach involving mass media campaigns, primary health-care interventions and advocacy is needed to draw attention to chronic kidney

disease and shift the focus from kidney failure to achieve a public health perspective on this disease.

- Annemiek Dorgelo
- Tom A. J. Oostrom

Comment 27 Jan 2022 Nature Reviews Nephrology



New insights into cognitive decline in chronic kidney disease

Chronic kidney disease (CKD) is known to be associated with cognitive impairment, but the mechanisms that underlie this kidney–brain connection are unclear. A recent study provides evidence that CKD is an independent risk factor for cognitive decline due to cerebral small vessel disease.

- Wei Ling Lau
- Mark Fisher

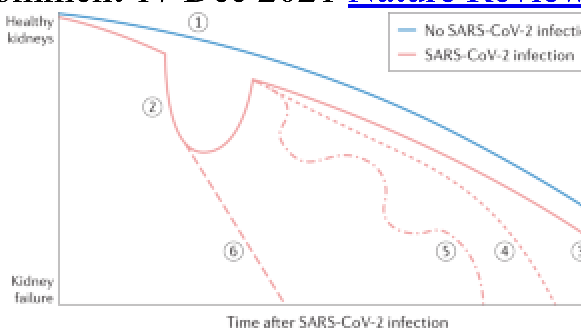
News & Views 2 Dec 2022 Nature Reviews Nephrology

Educating primary healthcare providers about kidney disease

Chronic kidney disease is a major public health problem that is associated with excessive morbidity, mortality and healthcare costs. However, limited clinician awareness of chronic kidney disease is universally identified as a key barrier to care. A concerted effort is urgently needed to address the knowledge gaps of primary care providers.

- Aminu K. Bello
- David W. Johnson

Comment 17 Dec 2021 Nature Reviews Nephrology

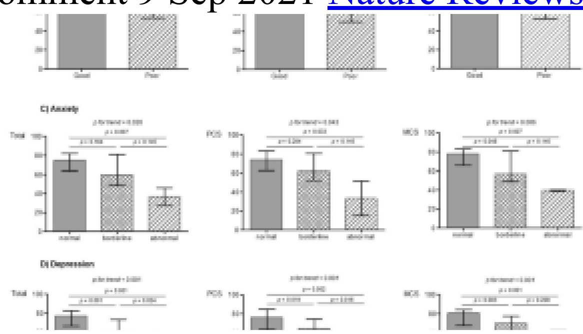


Long COVID and kidney disease

Kidney involvement is common in patients with acute SARS-CoV-2 infection, and subclinical inflammation and injury may persist for many months, resulting in a progressive decline in kidney function that leads to chronic kidney disease. Continued research is imperative to understand these long-term sequelae and identify interventions to mitigate them.

- Sachin Yende
- Chirag R. Parikh

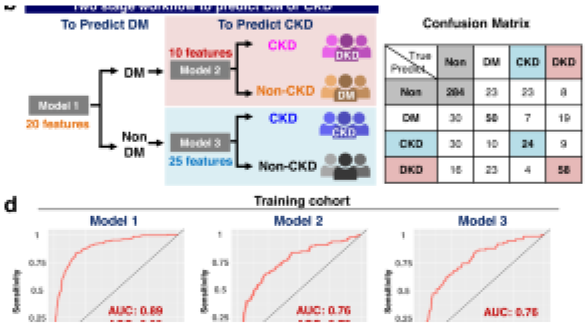
Comment 9 Sep 2021 Nature Reviews Nephrology



Depression is a main determinant of health-related quality of life in patients with diabetic kidney disease

- Suhyun Kim
- Junseok Jeon
- Jung Eun Lee

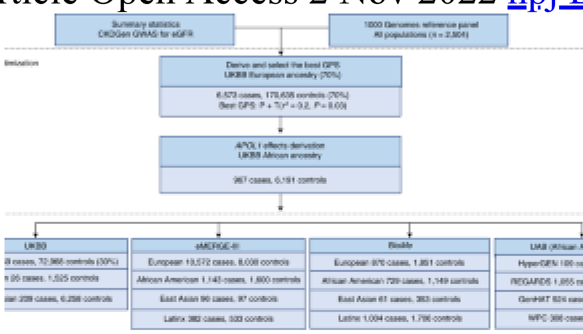
Article Open Access 16 Jul 2022 Scientific Reports SpringerLink



[Discovering a trans-omics biomarker signature that predisposes high risk diabetic patients to diabetic kidney disease]

- I-Wen Wu
- Tsung-Hsien Tsai
- Ting-Fen Tsai

Article Open Access 2 Nov 2022 npj Digital Medicine

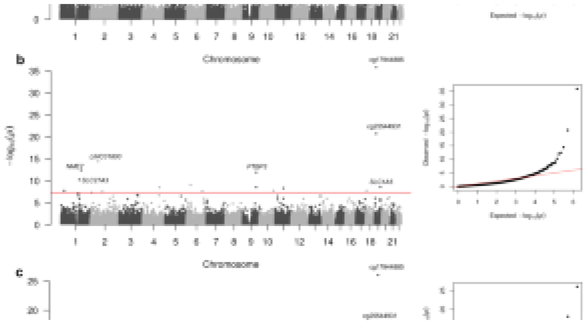


[Genome-wide polygenic score to predict chronic kidney disease across ancestries]

A new study generated and optimized a polygenic score for chronic kidney disease with reproducible performance across 15 cohorts of different ancestries, and identified potentially clinically relevant thresholds with predicted effects comparable to having a family history of the disease.

- Atlas Khan
- Michael C. Turchin
- Krzysztof Kiryluk

Article 16 Jun 2022 Nature Medicine SpringerLink

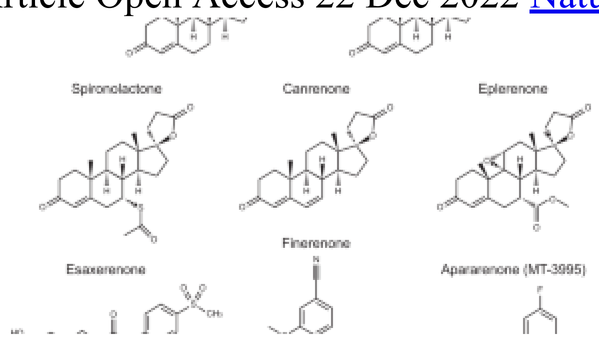


Epigenome-wide meta-analysis identifies DNA methylation biomarkers associated with diabetic kidney disease

Approximately 40 percent of people with type 1 diabetes develop kidney disease, but the risk factors are not well understood. Here, the authors identify DNA methylation signatures associated with diabetic kidney disease, of which 21 biomarkers predict the development of kidney failure.

- Laura J. Smyth
- Emma H. Dahlström
- Amy Jayne McKnight

Article Open Access 22 Dec 2022 Nature Communications



Clinical effect of nonsteroidal mineralocorticoid receptor (MR) antagonists in the treatment of diabetic kidney disease: expectations as a new therapeutic strategy

- Atsuhisa Sato
- Mitsuhiro Nishimoto

Review Article 20 Jun 2022 Hypertension Research

This article was downloaded by **calibre** from https://www.nature.com/collections/jbdjdcgjah



Is Japan’s research decline turning a corner?

Efforts on funding and helping early-career researchers could be bearing fruit, but major challenges remain.

- Benjamin Plackett

Nature Index 8 Mar 2023 Nature

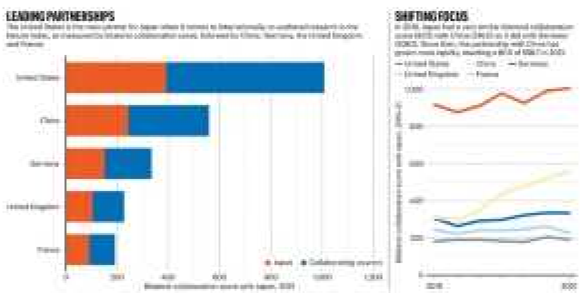


How Japanese science is trying to reassert its research strength

Successes in life sciences and international collaboration could be key to boosting the country’s high-quality output.

- Tim Hornyak

Nature Index 8 Mar 2023 Nature



Japan’s leading international partnerships

The country’s cross-border collaborations are varying across time and subject area.

Nature Index 8 Mar 2023 Nature



Freeing up Japan’s PhD potential

Better prospects are needed in universities and industry to make the most of valuable talent.

- Ayuko Hoshino

Nature Index 8 Mar 2023 Nature



Japan’s rising research stars: Mariko Kimura

Helped by amateur astronomers, Kimura is delving into the mysteries of distant galaxies.

- Chris Woolston

Nature Index 8 Mar 2023 Nature



Japan’s rising research stars: Yuuki Wada

Wada is studying high-energy particles by collecting data from storms on Earth.

- Chris Woolston

Nature Index 8 Mar 2023 Nature

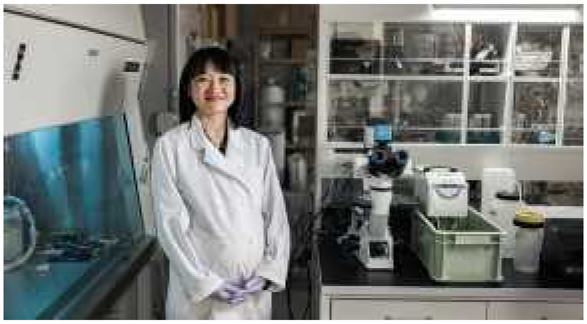


Japan’s rising research stars: Tatsuya Kubota

Using submarine sensors, Kubota works to understand earthquakes and tsunamis.

- Sandy Ong

Nature Index 8 Mar 2023 Nature



Japan’s rising research stars: Yasuka Toda

Toda’s research shows that a love for umami allowed our evolutionary ancestors to start eating plants.

- Sandy Ong

Nature Index 8 Mar 2023 Nature



Japan’s rising research stars: Ken-ichi Otake

Otake uses novel materials to isolate ‘heavy water’ cheaply and efficiently.

- Sandy Ong

Nature Index 8 Mar 2023 Nature



[Will Japan’s new ¥10-trillion university fund lift research performance?]

Endowment inspired by US Ivy League model must be accompanied by other reforms, critics say.

- David McNeill

Nature Index 8 Mar 2023 Nature

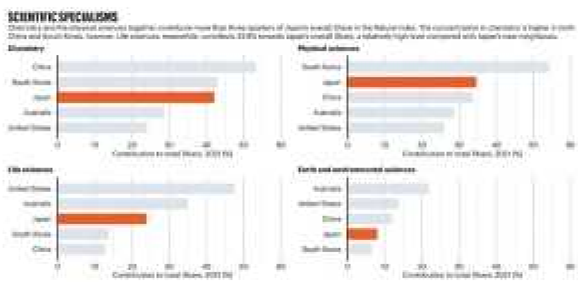


Shoring up Japan’s research performance

The government’s university endowment scheme is a ‘big moment’ for Japanese science, says Takahiro Ueyama.

- David McNeill

Nature Index 8 Mar 2023 Nature



[How Japan's disciplinary strengths are shifting focus]

Physical sciences and chemistry still dominate research output, but life sciences is increasingly important.

Nature Index 8 Mar 2023 Nature



Japanese robotics lags as AI captures global attention

The country's automation research might need a renewed focus to rekindle past successes.

- Rachel Nuwer

Nature Index 8 Mar 2023 Nature



A guide to the Nature Index

A description of the terminology and methodology used in this supplement, and a guide to the functionality that is available free online at natureindex.com.

Nature Index 8 Mar 2023 Nature

Leading 100 Japanese institutions

Data Page 8 Mar 2023

Leading 25 Japanese institutions in chemistry

Data Page 8 Mar 2023

Top 10 rising Japanese institutions

Data Page 8 Mar 2023

Leading 25 Japanese institutions in Earth & environmental sciences

Data Page 8 Mar 2023

Leading 25 Japanese institutions in life sciences

Data Page 8 Mar 2023

Leading 25 Japanese institutions in physical sciences

Data Page 8 Mar 2023



[Mapping aquatic biodiversity and pollution in detail]

Researchers are monitoring multiple underwater species dynamics for the first time using tiny snippets of shed environmental DNA, while others are tracing the complex routes of mineral pollutants.

Advertisement Feature 8 Mar 2023



[The science and art of challenging assumptions]

Some of the biggest discoveries in history have been made when researchers challenged the status-quo – a lesson two Tokyo-based academics have taken to heart.

Advertisement Feature 22 Feb 2023



How industrial exhaust heat could be used to extract carbon dioxide

Clever use of absorbent materials could help harness the 80% of industrial exhaust heat energy that is currently wasted to power carbon capture technology.

Advertisement Feature 8 Mar 2023



Revealing the secrets of metabolism

Understanding how cells use glucose will provide insights into how they work.

Advertisement Feature 8 Mar 2023



Trial shows the best way to treat colorectal cancer

A large-scale Japanese trial provides vital evidence for the most effective antibody for treating certain inoperable colorectal cancers.

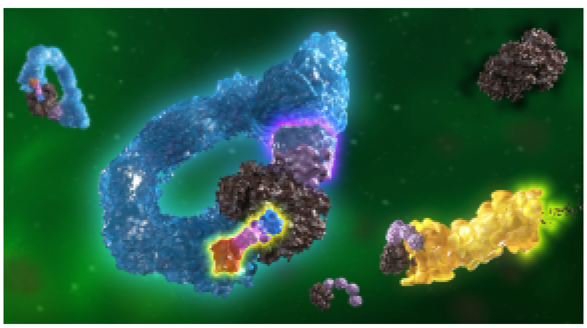
Advertisement Feature 8 Mar 2023



How X-rays are empowering better batteries

New X-ray diffraction tools are supercharging the search for superior lithium-ion battery materials.

Advertisement Feature 8 Mar 2023



Novel cancer-drug strategy could hit notorious proteins

Undruggable targets are proving no match for emerging protein-degrader technologies.

Advertisement Feature 8 Mar 2023



Cell culture medium is key to regenerative medicine

A Japanese pharmaceutical company has set its sights on advancing the field of regenerative medicine, by developing a new liquid medium for growing human stem cells best suited for stem-cell therapy.

Advertisement Feature 8 Mar 2023



[The light-based tech behind some of physics’ biggest breakthroughs]

From sensing cosmic neutrinos to detecting the elusive Higgs boson, light-based technology is empowering breakthroughs in physics.

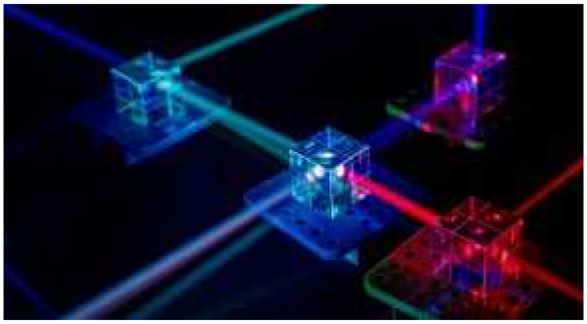
Advertisement Feature 8 Mar 2023



[Harnessing AI to build a fresh vision for creativity]

Artificial-intelligence research is seeking to enhance communication between humans and machines and find new ways for AI to contribute to society.

Advertisement Feature 8 Mar 2023



['Magic wands' for laser tech]

Versatile optical crystals are benefiting telecommunications and advancing quantum computing.

Advertisement Feature 8 Mar 2023



[Supporting depression diagnosis through machine learning]

A machine-learning algorithm capable of identifying patients with major depressive disorder could transform diagnostics and treatments.

Advertisement Feature 8 Mar 2023



[Cryo-electron microscopy offers cool potential]

The visualization of proteins in unprecedented detail is spurring drug discovery and could help to treat rare cancers.

Advertisement Feature 8 Mar 2023



Japanese lacquer’s natural advantages

A polymer chemist is exploring eco-friendly, gloss paint alternatives and is looking to Japan’s 9000-year-old tradition of lacquering for inspiration.

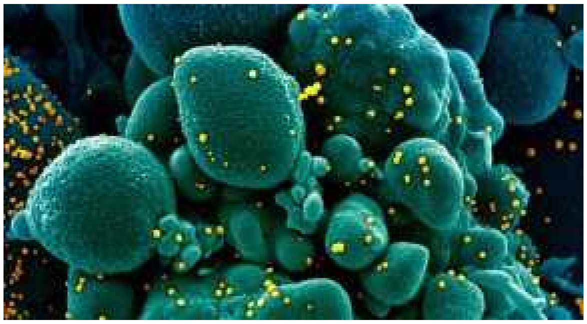
Advertisement Feature 8 Mar 2023



Deepening understanding of human civilization

Researchers in Japan are adopting a novel multi-disciplinary approach to uncover lessons from the past in both Asia and Mesoamerica.

Advertisement Feature 8 Mar 2023



Taking great technological strides for medical science

COVID-19 sparked intense interest in vaccines. Some Japanese immunologists took up the gauntlet to help slow the spread of the virus, while others have strived to digitally map human biology.

Advertisement Feature 8 Mar 2023



Powerful lasers herald age of new human adventures

Hugely powerful lasers are shining light on the cosmos and can generate energy using the same process that occurs in the scorching interior of stars.

Advertisement Feature 8 Mar 2023



The surprising new ideas coming from AI

From powering miniature lab-based brains, to helping humans understand quantum mechanics, cross-pollinating AI with other disciplines is creating new and surprising research directions.

Advertisement Feature 8 Mar 2023

This article was downloaded by **calibre** from https://www.nature.com/collections/ahdgheahhc